WIND-TUNNEL VALIDATION OF A REINFORCEMENT LEARNING BASED GUST LOAD ALLEVIATION CONTROLLER FOR A FLEXIBLE WING

Ramesh Konatala[1], **Felix Stalla**[1], **Gertjan Looye**[1], **Thomas G. Schmidt**[2], **Lukas Koida**[2], **Charlotte Hanke**[2], **Erik-Jan van Kampen**[3]

[1]German Aerospace Center (DLR), Institute of Aeroelasticity, Control of Aeroelastic Systems
82234 Wessling, Germany
(corresponding author e-mail: ramesh.konatala@dlr.de)

[2]German Aerospace Center (DLR), Institute of Aeroelasticity, Aeroelastic Experiments
37073 Goettingen, Germany

[3]Delft University of Technology
Kluyverweg 1, 2629 HS Delft, The Netherlands

Keywords: reinforcement learning, soft actor–critic, gust load alleviation, aeroelastic control, flexible wing, wind-tunnel testing

Abstract: This paper addresses the experimental validation of a reinforcement learning (RL)-based gust load alleviation (GLA) controller for a flexible aeroelastic wing. A soft actor–critic (SAC) controller is trained offline using a linear aeroservoelastic (ASE) model at 30 m/s with broadband chirp gust excitation and is deployed as a fixed deterministic neural-network policy. The policy maps outboard acceleration and flap-deflection measurements to trailing-edge flap commands in real time. Before wind-tunnel testing, the trained policy is evaluated in simulation using experimentally identified actuator and sensor effects, including delay, position and rate limits, backlash, and measurement noise. Wind-tunnel experiments are performed in the DNW–NWB wind tunnel under harmonic gust excitation generated by rotating slotted-cylinder gust generators. For the single-actuator Flap 4 configuration, the controller reduces the peak-to-peak wing root bending moment (WRBM) by approximately 65% at the first bending-mode frequency. When the gust excitation is increased using two generators, the same fixed Flap 4 policy maintains a 51% reduction. These results show that an SAC policy trained on a linear ASE model can be transferred to wind-tunnel hardware for targeted GLA when supported by simulation-based verification and validation (V&V).

1 INTRODUCTION

The pursuit of improved fuel efficiency in modern aircraft has motivated the development of lightweight airframes, advanced composite structures, and high-aspect-ratio wings. While these design trends reduce induced drag and structural weight, they also increase structural flexibility and strengthen the coupling between aerodynamic, elastic, and inertial forces. As a result, a given gust disturbance can produce larger flexible structural responses and higher dynamic loads, making the control of aeroelastic dynamics increasingly relevant for future aircraft design [1–3]. Gust-induced loads must remain within prescribed certification limits, as reflected in regulatory requirements for discrete gusts and continuous turbulence [4, 5]. Active aeroelastic control therefore plays an enabling role by using distributed sensing, high-bandwidth actuation, and feedback control to modify the structural response in real time. Within this broad class of methods, gust load alleviation (GLA) aims to reduce gust-induced structural loads, whereas active flutter suppression (AFS) aims to increase modal damping and delay flutter onset.

Active aeroelastic control has traditionally relied on model-based control synthesis. In this workflow, a high-fidelity aeroservoelastic (ASE) representation is assembled from structural, aerodynamic, actuator, and sensor models, reduced to retain the dominant flexible modes, and linearized around selected operating conditions. The resulting linear ASE models are then used for robust or optimal control synthesis. This approach has led to successful experimental demonstrations, including active control of an experimental flexible wing [6], wind-tunnel testing of active gust load alleviation [7], and flight testing of blending-based active flutter suppression [8]. However, the same workflow also introduces practical limitations. Accurate ASE modelling and validation can be time consuming, and model reduction and linearization may neglect nonlinear actuator, sensor, aerodynamic, or structural effects that become relevant during implementation. Moreover, flexible-wing configurations with multiple sensors and control effectors can make multivariable controller synthesis, tuning, and scheduling increasingly complex. For GLA, these challenges are amplified by high-order aeroelastic dynamics, competing time- and frequency-domain objectives, actuator constraints, and robustness requirements [9].

These challenges motivate the investigation of data-driven and learning-based control methods for aeroelastic applications. A neural-network controller is attractive because it can represent nonlinear sensor-to-actuator mappings beyond a prescribed linear feedback structure. Multilayer feedforward networks are known to be universal approximators, indicating that, with sufficient capacity, they can represent broad classes of nonlinear mappings [10]. However, representation capability alone does not determine suitable controller parameters, nor does it guarantee useful behavior outside the finite training experience. This distinction is central in learning theory, where the generalization of learned mappings from finite examples is explicitly studied [11]. reinforcement learning (RL) addresses the parameter-search problem by updating a policy through interaction with an environment and a reward function that encodes the control objective. Among RL methods, soft actor–critic (SAC) is particularly relevant for continuous-control problems because it combines an off-policy actor–critic structure with maximum-entropy exploration [12]. For aeroelastic control, this formulation is attractive because the same training framework can, in principle, use linear models, nonlinear simulation environments, or plant-derived data.

Recent studies already indicate the potential of learning-based control for aeroelastic gust-response problems. Haughn et al. demonstrated deep-learning-based gust rejection on a smart-material camber-morphing wing, showing that gust impact could be reduced using real-time onboard pressure measurements while reducing sensor requirements [13]. Hu et al. investigated RL-based gust load control of an elastic wing with camber morphing, where a policy trained using a simplified aeroelastic model was assessed using high-fidelity CFD-based fluid–structure-control simulations [14]. However, experimental validation of fixed neural-network GLA policies on flexible aeroelastic wind-tunnel hardware remains limited. Building on this emerging direction, the present paper demonstrates SAC-based GLA on a flexible aeroelastic wind-tunnel demonstrator equipped with discrete trailing-edge flaps. The scope is deliberately limited to an offline-trained controller using a linear ASE training model. Before wind-tunnel testing, the trained controller is evaluated using a separate simulation model that includes experimentally identified nonlinear actuator and sensor dynamics. The main contribution is a wind-tunnel validation of a fixed SAC neural-network policy for GLA, with emphasis on load reduction, actuator behavior, repeatability, gust-intensity sensitivity, and sim-to-real agreement.

The remainder of this paper is organized as follows. Section 2 introduces the flexible-wing aeroelastic demonstrator modelling workflow, including the structural, aerodynamic, gust, actuator, and sensor models used to construct the linear ASE training environment and the nonlinear actuator–sensor simulation model used before wind-tunnel testing. Section 3 presents the SAC-based GLA control design, including the observation and action definitions, actor–critic neural-network architecture, reward formulation, and offline training procedure. Section 4 describes the experimental setup, simulation analyses, and wind-tunnel validation results, with emphasis on wing root bending moment (WRBM) attenuation, actuator-configuration effects, repeatability, gust-intensity sensitivity, and actuator-rate limitations. Finally, Section 5 summarizes the main findings and discusses future extensions toward nonlinear-model training, offline fine-tuning with plant-derived data, in-flight validation, and broader learning-based active aeroelastic control applications.

2 AEROSERVOELASTIC MODELING OF THE FLEXIBLE WING

Although the SAC controller does not require an explicit model during deployment, a representative ASE model is required for two purposes: to generate realistic interaction data during offline training and to assess the trained policy before wind-tunnel testing. The modeling procedure follows the modular VarLoads framework and the unified manoeuvre/gust-load formulation described in [15, 16]. The model consists of a reduced structural model, an unsteady aerodynamic model, a gust model, actuator and sensor models, and wing-root load recovery. Together, these components define the gust-to-load, gust-to-sensor, and actuator-to-load dynamics relevant for the GLA task. The resulting linear state-space model is used for SAC training, while a nonlinear model with identified actuator and sensor effects is used for closed-loop simulation verification.

2.1 Structural dynamics

The flexible wing structure is represented by a finite-element model that is statically condensed and reduced to a modal basis. Since the wind-tunnel model is clamped at the root, rigid-body flight dynamics are not retained. The structural dynamics are therefore expressed only in terms of the flexible modal coordinates. Following the mean-axis formulation used in loads analysis [16, 17], the modally reduced equation of motion is

$$\mathbf{M}_{ff}\ddot{\mathbf{u}}_f + \mathbf{B}_{ff}\dot{\mathbf{u}}_f + \mathbf{K}_{ff}\mathbf{u}_f = \mathbf{\Phi}_{gf}^T\mathbf{P}_g^{ext}. \tag{1}$$

Here, $\mathbf{u}_f \in \mathbb{R}^{15}$ denotes the retained flexible modal displacement vector, while $\dot{\mathbf{u}}_f$ and $\ddot{\mathbf{u}}_f$ denote modal velocity and acceleration. The matrices $\mathbf{M}_{ff}$, $\mathbf{B}_{ff}$, and $\mathbf{K}_{ff}$ are the modal mass, damping, and stiffness matrices, respectively. The matrix $\mathbf{\Phi}_{gf}$ maps the retained flexible modes to the structural grid, denoted by the g-set, and $\mathbf{P}_g^{ext}$ contains the external forces and moments acting on the structural grid. In the present wind-tunnel configuration, these external loads are dominated by aerodynamic and gust-induced forces.

2.2 Unsteady aerodynamic model

The unsteady aerodynamic loads are computed using the Doublet Lattice Method (DLM) on the discretized aerodynamic lifting surface. In the frequency domain, the aerodynamic forces and moments acting on the structural grid are written as

$$\mathbf{P}_g^{aero}(k) = q_\infty \mathbf{T}_{kg}^T \mathbf{S}_{kj} \mathbf{Q}_{jj}(k)\mathbf{w}_j, \qquad k = \frac{\omega c_{\mathrm{ref}}}{2U_\infty}. \tag{2}$$

In Eq. (2), q_∞ is the dynamic pressure, U_∞ is the freestream velocity, c_{ref} is the aerodynamic reference chord, and k is the reduced frequency. The matrix $\mathbf{Q}_{jj}(k)$ is the aerodynamic influence coefficient matrix, and $\mathbf{w}_j$ is the downwash vector at the aerodynamic control points. The matrix $\mathbf{S}_{kj}$ converts pressure coefficients into aerodynamic forces and moments at the aerodynamic grid, while $\mathbf{T}_{kg}$ couples the aerodynamic grid to the structural grid using spline-based aero-structural interpolation.

The pressure jump on the aerodynamic panels is therefore

$$\Delta\mathbf{c}_{p,j}(k) = \mathbf{Q}_{jj}(k)\mathbf{w}_j(k). \tag{3}$$

The downwash is generated by the flexible motion of the wing, by control-surface motion, and by the external gust disturbance. In compact form,

$$\mathbf{w}_j = \mathbf{w}_j^f + \mathbf{w}_j^{x2} + \mathbf{w}_j^G, \tag{4}$$

where $\mathbf{w}_j^f$ is the contribution from the modal displacement and velocity, $\mathbf{w}_j^{x2}$ is the contribution from the trailing-edge flap deflections and their derivatives, and $\mathbf{w}_j^G$ is the gust-induced downwash.

Since the DLM provides aerodynamic forces in the frequency domain, a time-domain realization is required for simulation and training. The unsteady aerodynamic influence matrices are approximated using Roger's rational function approximation (RFA). With

$$p = \frac{s c_{\mathrm{ref}}}{2U_\infty}, \tag{5}$$

the aerodynamic influence matrix is approximated as

$$\mathbf{Q}_{jj}(p) \approx \mathbf{A}_0 + p\mathbf{A}_1 + p^2\mathbf{A}_2 + \sum_{\ell=1}^{n_\ell} \mathbf{A}_{\ell+2}\frac{p}{p+\beta_\ell}. \tag{6}$$

The matrices $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_{\ell+2}$ are RFA coefficient matrices, and β_ℓ are the selected aerodynamic lag poles. This approximation introduces aerodynamic lag states and enables the DLM-based aerodynamic model to be integrated with the structural dynamics in the time domain.

2.3 Gust disturbance model

The gust disturbance is introduced through the gust angle,

$$w_G(t) \equiv \alpha_G(t) = \frac{u_G(t)}{U_\infty}, \tag{7}$$

where $u_G(t)$ is the vertical gust velocity and U_∞ is the freestream velocity. The gust model maps this scalar gust-angle input to an equivalent aerodynamic downwash distribution and then to external aerodynamic forces on the wing. In state-space form, the gust-aerodynamic contribution can be represented as

$$\dot{\mathbf{x}}_G = \mathbf{A}_G\mathbf{x}_G + \mathbf{B}_G w_G, \tag{8}$$

$$\mathbf{w}_j^G = \mathbf{C}_G\mathbf{x}_G + \mathbf{D}_G w_G. \tag{9}$$

The gust states $\mathbf{x}_G$ account for the penetration and unsteady aerodynamic effects associated with the gust field. For discrete gust simulations, a one-minus-cosine gust profile can be expressed as

$$w_G(t) = \begin{cases} \hat{\alpha}_G \left[1 - \cos\left(\dfrac{2\pi U_\infty t}{2H}\right)\right], & t_s \le t \le t_e, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

where $\hat{\alpha}_G$ is the gust-angle parameter, H is the gust gradient, and t_s and t_e denote the start and end times of the gust encounter. For harmonic gust excitation, the same input channel w_G is used with a sinusoidal gust-angle signal.

2.4 Linear aeroservoelastic realization

Combining the structural dynamics, unsteady aerodynamics, control-surface aerodynamics, and gust model yields the linear aeroelastic realization

$$\begin{aligned} \dot{\mathbf{x}}_{AE} &= \mathbf{A}_{AE}\mathbf{x}_{AE} + \mathbf{B}_{AE}\mathbf{v}_{AE}, & \mathbf{y}_{AE} &= \mathbf{C}_{AE}\mathbf{x}_{AE} + \mathbf{D}_{AE}\mathbf{v}_{AE}, \\ \mathbf{v}_{AE} &= \begin{bmatrix} w_G & \mathbf{u}^T & \dot{\mathbf{u}}^T & \ddot{\mathbf{u}}^T \end{bmatrix}^T, & \mathbf{x}_{AE} &= \begin{bmatrix} \mathbf{u}_f^T & \dot{\mathbf{u}}_f^T & \mathbf{x}_{f,\text{lag}}^T & \mathbf{x}_{u,\text{lag}}^T & \mathbf{x}_G^T \end{bmatrix}^T. \end{aligned} \tag{11}$$

Here, $\mathbf{u} = [u_1, u_2, u_3, u_4]^T$ contains the four trailing-edge flap deflections, with Flap 4 located closest to the wing tip. All flap deflections are expressed in radians, and positive deflection corresponds to trailing-edge down. The aeroelastic model contains 229 states: 30 structural modal states, 120 flexible-motion aerodynamic lag states, 32 control-surface aerodynamic lag states, and 47 gust-model states. The output vector contains the modal displacements, velocities, accelerations, and six wing-root balance loads, $\mathbf{P}_{WR} = [F_x, F_y, F_z, M_x, M_y, M_z]^T$. In the adopted wind-tunnel balance convention, M_x is the WRBM, and M_y is the wing root torsion moment (WRTM).

The aeroelastic model is augmented with a first-order linear actuator model for each flap,

$$\frac{u_i(s)}{u_{\text{cmd},i}(s)} = \frac{\omega_{act}}{s + \omega_{act}}, \qquad i = 1, \ldots, 4, \tag{12}$$

where $u_{\mathrm{cmd},i}$ is the commanded flap deflection. The actuator states are the realized flap deflections, while the corresponding flap rates and accelerations are supplied to the aeroelastic model through $\mathbf{v}_{AE}$. The resulting linear ASE model is written compactly as

$$\dot{\mathbf{x}}_{ASE} = \mathbf{A}_{ASE}\mathbf{x}_{ASE} + \mathbf{B}_G w_G + \mathbf{B}_u \mathbf{u}_{\mathrm{cmd}}, \qquad \mathbf{y}_{ASE} = \mathbf{C}_{ASE}\mathbf{x}_{ASE} + \mathbf{D}_G w_G + \mathbf{D}_u \mathbf{u}_{\mathrm{cmd}}, \quad (13)$$

with $\mathbf{x}_{ASE} = [\mathbf{x}_{AE}^T, \mathbf{x}_{act}^T]^T \in \mathbb{R}^{233}$. Thus, the four actuator states increase the model order from 229 aeroelastic states to 233 ASE states.

The acceleration sensor model is a static transformation from modal accelerations to physical accelerations at the sensor locations. The full sensor model maps the 15 modal accelerations to ten vertical accelerometer channels, denoted by stations $1a$–$5b$. The labels a and b indicate sensors located closer to the leading and trailing edge, respectively. For SAC training and wind-tunnel testing, only the four outboard acceleration measurements are used,

$$\ddot{\mathbf{u}}_m = \mathbf{T}_{sens}\ddot{\mathbf{u}}_f, \qquad \mathbf{a}_{obs} = \begin{bmatrix} \ddot{u}_{m,4a} & \ddot{u}_{m,4b} & \ddot{u}_{m,5a} & \ddot{u}_{m,5b} \end{bmatrix}^T. \quad (14)$$

The root loads are recovered using the force summation method, where the inertial loads are subtracted from the external loads on the structural grid and then summed to obtain the six-axis wing-root balance loads,

$$\mathbf{P}_g^{load} = \mathbf{P}_g^{ext} - \mathbf{M}_{gg}\mathbf{\Phi}_{gf}\ddot{\mathbf{u}}_f, \qquad \mathbf{P}_{WR} = \mathbf{R}_{WR}\mathbf{P}_g^{load}. \quad (15)$$

The linear ASE model used for SAC training is trimmed at $U_\infty = 30$ m/s and $\alpha = 0°$, and discretized with a sampling time of $T_s = 1$ ms, consistent with the real-time control implementation. Additional models trimmed at 40 m/s and 50 m/s are used only for desktop simulation checks to assess sensitivity to operating-speed variations.

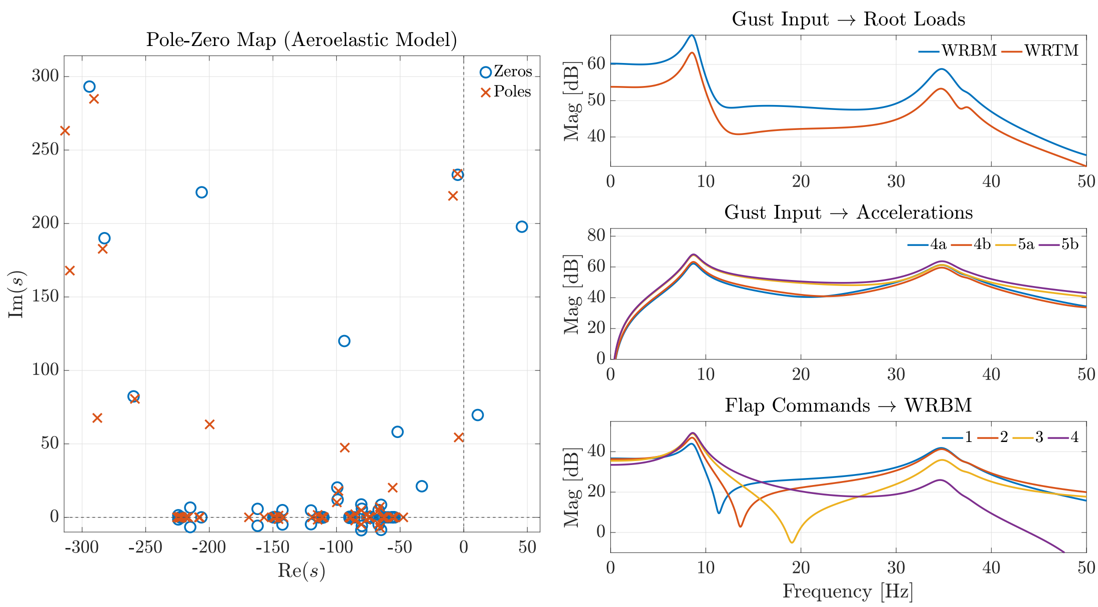


Figure 1: Pole–zero map and frequency-response characteristics of the aeroelastic model used for SAC-based GLA training. The gust-to-load transfer functions identify the dominant flexible modes, the gust-to-acceleration responses indicate the observability of the outboard sensor set, and the flap-to-WRBM transfer functions quantify the relative control authority of the four trailing-edge flaps.

Figure 1 summarizes the aeroelastic characteristics relevant to the GLA control task. The pole–zero map identifies the dominant flexible modes within the investigated frequency range.

The gust-to-load transfer functions show that the first wing bending mode dominates the gust-induced WRBM response, while the gust-to-acceleration responses confirm that the selected outboard accelerometers observe the relevant flexible dynamics. The flap-to-WRBM transfer functions indicate the relative load-alleviation authority of the four trailing-edge flaps and support the selection of the outboard flap as the baseline single-actuator configuration.

2.5 Nonlinear actuator and sensor model

A separate nonlinear simulation model is used for implementation-aware closed-loop verification before wind-tunnel testing. This model preserves the linear ASE structure in Eq. (13), but replaces the ideal actuator and sensor blocks with experimentally identified dynamics summarized in Table 1. These effects are excluded from SAC training and are reserved for closed-loop simulation analyses and comparison with wind-tunnel measurements.

Table 1: Nonlinear Actuator and Sensor Dynamics

Component	**Parameter**	**Value**
Digital Control	Sampling time (T_s)	1 ms
Actuator Model	First-order time constant (τ)	0.011 s
	Time delay (deadtime)	6 ms
	Position limit	$\pm 16°$
	Rate limit	$\pm 700°$/s
	Deadband (backlash)	1°
Accelerometers	Measurement delay	3 ms
	Gaussian noise (1σ)	1.0 m/s^2

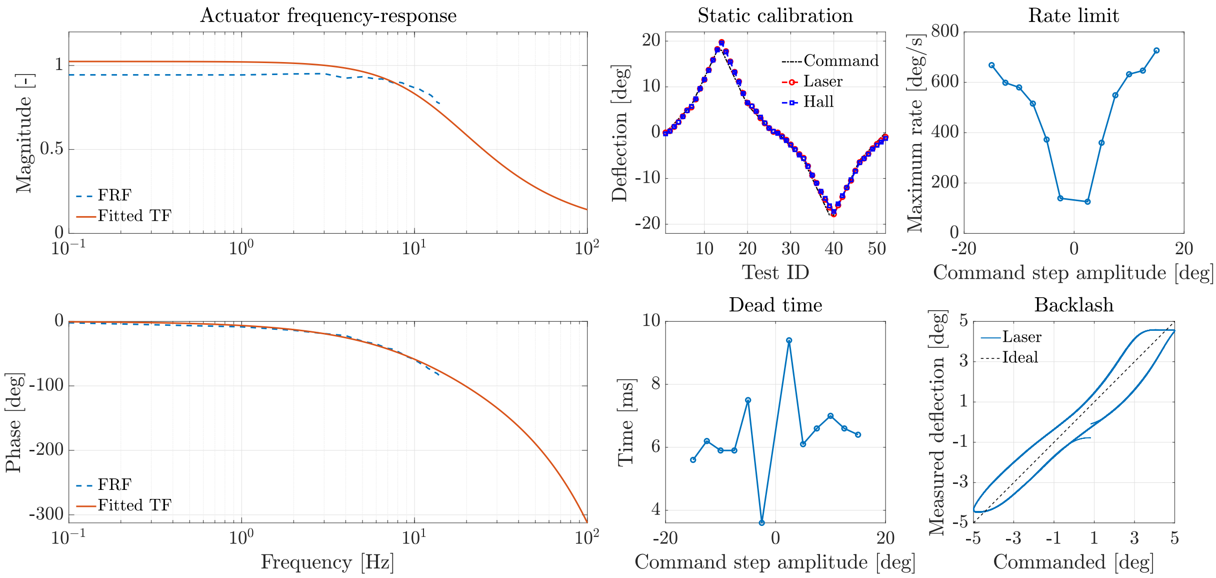


Figure 2: Actuator characterization for Flap 4. The frequency-response data are used to fit the first-order actuator dynamics, while static calibration, step-response, and low-amplitude cyclic tests identify the deflection scaling, position and rate limits, dead time, and backlash.

Figure 2 shows the identification procedure for Flap 4 as a representative example. Following the black-box actuator identification approach of [18], the assembled command–actuator–linkage–flap system is characterized experimentally rather than modelled from its internal servo dynamics. Sinusoidal excitations at different frequencies are first used to obtain the command-to-measured deflection frequency response function (FRF), from which the first-order actuator time constant is fitted. Static calibration tests compare commanded deflection with laser-triangulation and Hall-sensor measurements, thereby establishing the deflection scaling, Hall-sensor consistency, and usable position range. Step-response tests with different command amplitudes are then used to estimate the achievable flap-rate limit from the maximum slope of the measured deflection and the command-to-response dead time from the onset of flap motion. Finally, low-amplitude harmonic excitation is used to quantify backlash from the hysteresis between commanded and measured deflection; this effect is represented using a Simulink backlash block in the nonlinear actuator model. The same procedure is applied to the remaining control surfaces, and the identified nonlinear actuator and sensor models are integrated into the closed-loop verification environment.

3 SOFT ACTOR–CRITIC-BASED GUST LOAD ALLEVIATION CONTROL DESIGN

The objective of the GLA control system is to command the trailing-edge flaps such that gust-induced structural loads on the flexible wing are reduced. In this work, the WRBM is used as the primary performance metric for evaluating load alleviation, but it is not used as a feedback signal or as a direct reward term during training. Instead, the SAC agent is trained using local acceleration feedback from wing-mounted sensors, which provides an experimentally available proxy for the gust-induced flexible response. The controller learns a direct mapping from measured sensor signals to actuator commands through RL. The SAC algorithm [12] is an off-policy actor–critic method formulated within the maximum-entropy framework. In contrast to entropy-unregularized actor–critic formulations, SAC augments the return objective with a policy-entropy term to promote exploration and improve training robustness. The sensor–actuator configuration and control architecture build on previous wind-tunnel studies of active GLA control on the same flexible-wing demonstrator [7], where distributed trailing-edge flaps and outboard acceleration feedback were shown to be effective for mitigating gust-induced loads. The present work focuses mainly on the single-actuator baseline controller using Flap 4. Nevertheless, the same training framework was also applied to two-, three-, and four-actuator configurations by extending the observation and action vectors accordingly.

3.1 Aeroelastic control law architecture

Figure 3 illustrates the SAC-based GLA control law architecture. During offline training, the SAC agent consists of a stochastic actor and two Q-value critics. The actor maps normalized sensor feedback to continuous flap commands, while the critics are used only during training to estimate the state–action value functions. After training, only the actor is retained for closed-loop evaluation and real-time deployment.

For the baseline single-actuator controller, the observation vector is defined as

$$\mathbf{s} = \begin{bmatrix} \ddot{x}_{4a} & \ddot{x}_{4b} & \ddot{x}_{5a} & \ddot{x}_{5b} & \delta_4 \end{bmatrix}^T, \tag{16}$$

where $\ddot{x}_{4a}$, $\ddot{x}_{4b}$, $\ddot{x}_{5a}$, and $\ddot{x}_{5b}$ are the vertical acceleration measurements from sensors located near Flaps 3 and 4, and δ_4 is the measured Flap 4 deflection obtained from the Hall sensor. The acceleration signals are expressed in SI units, while the measured flap deflection δ_4 included in the observation vector is expressed in radians. The actor output a is bounded $([-10^\circ, 10^\circ])$ as

$$a \in [-1, 1], \qquad \delta_{4,\mathrm{cmd}} = \delta_{\max} a, \qquad \delta_{\max} = 10^\circ = 0.17453\,\mathrm{rad}. \tag{17}$$

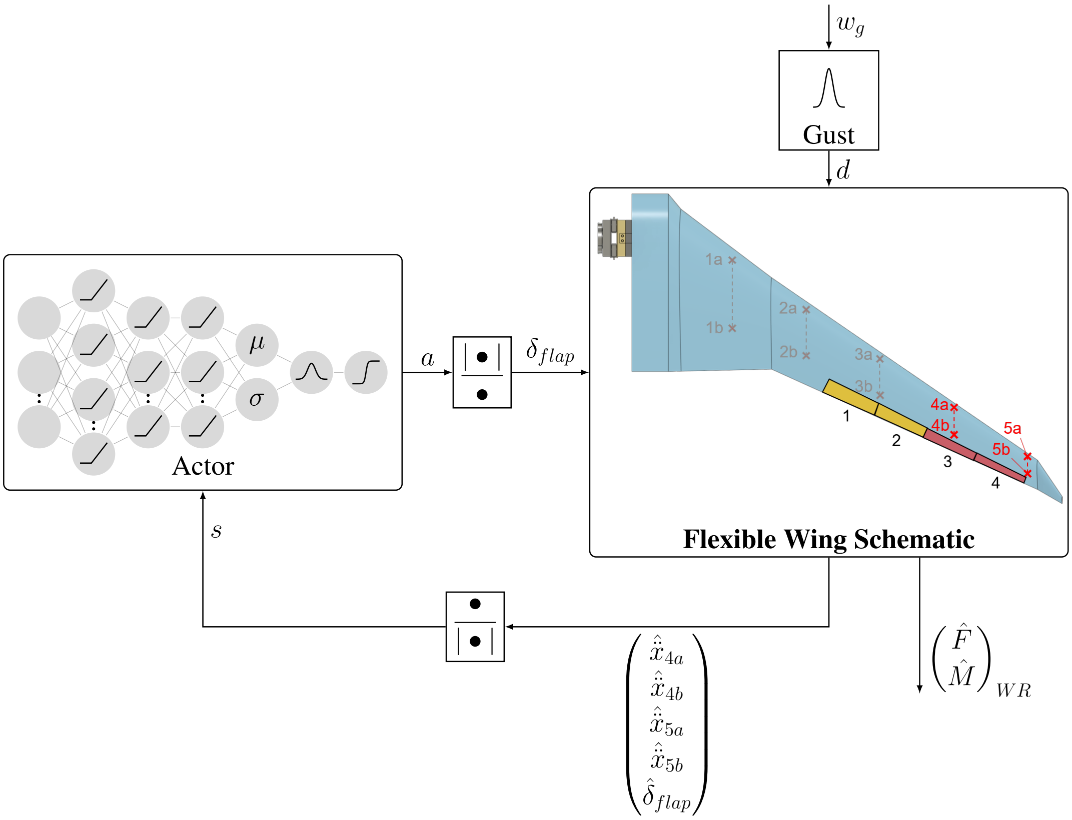


Figure 3: Architecture of the Soft Actor–Critic (SAC) based gust load alleviation (GLA) control law. The SAC agent, composed of an actor and two critics, interacts with the aeroelastic model of the flexible wing to suppress gust-induced loads using measured acceleration and flap deflection feedback signals. The gust block represents the aerodynamic gust model, representing external gust generators in wind tunnel experiments.

where $\delta_{4,\text{cmd}}$ is the commanded Flap 4 deflection and positive flap deflection corresponds to trailing-edge-down motion.

For a controller using n_a active flaps, the same structure is extended as

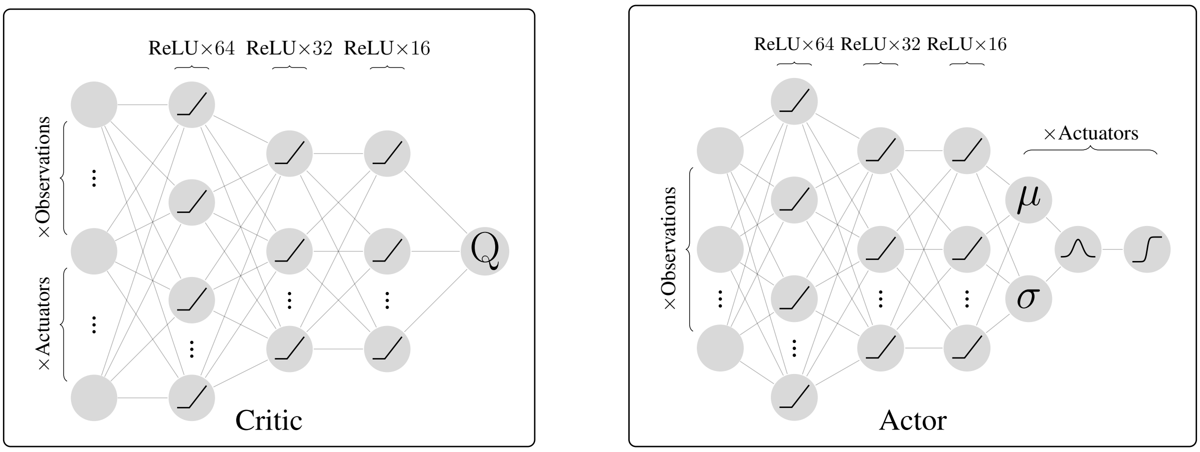
$$\mathbf{s} = \begin{bmatrix}\ddot{x}_{4a} & \ddot{x}_{4b} & \ddot{x}_{5a} & \ddot{x}_{5b} & \delta_{i_1} & \cdots & \delta_{i_{n_a}}\end{bmatrix}^T, \tag{18}$$

and

$$\mathbf{a} \in [-1, 1]^{n_a}, \qquad \delta_{i_j,\text{cmd}} = 10^\circ a_j, \quad j = 1, \ldots, n_a. \tag{19}$$

Thus, changing the actuator configuration mainly requires changing the number of measured flap-deflection feedback signals and the number of commanded flap outputs. This allows different actuator configurations to be studied by changing the observation and action dimensions and retraining the policy, rather than deriving a separate controller structure for each case.

During training, no additional actuator dynamics, rate limits, backlash, measurement filters, sensor noise, or sensor delays were included. The flap command was applied through a direct feedthrough path, and the only actuator constraint during training was the actor-output scaling to $\pm 10^\circ$. This choice was made to maximize exploration of the aeroelastic state space during learning. Actuator dynamics, deflection limits, rate limits, sensor filtering, and nonlinear actuator effects were introduced later during simulation-based verification and experimental deployment, as described in Section 4.



(a) Critic network architecture.

(b) Actor network architecture.

Figure 4: Neural network architectures used for the SAC controller. The critic network is used for both Q-functions, while the actor network maps the observation vector to the parameters of a Gaussian policy, followed by sampling and $\tanh$ squashing to obtain bounded flap commands.

3.2 SAC agent structure

Figure 4 shows the neural-network architectures used for the SAC controller. The actor implements a stochastic Gaussian policy $\pi(\mathbf{a}|\mathbf{s};\theta)$, which maps the normalized observation vector $\mathbf{s}$ to the parameters of a continuous action distribution. The actor contains fully connected layers with 64 and 32 neurons in the shared observation path. The mean branch contains an additional hidden layer with 16 neurons before the action-mean output, while the standard-deviation branch uses a softplus output to ensure positive standard deviations. ReLU activation functions are used in the hidden layers. For bounded continuous actions, the sampled Gaussian action is passed through a tanh operation and scaled to the action limits by the SAC agent implementation [19].

The SAC agent uses two Q-value critics, denoted by $Q_k(\mathbf{s},\mathbf{a};\phi_k)$, $k \in \{1,2\}$. Each critic receives the observation and action as inputs and estimates the corresponding state–action value. The use of two critics helps reduce overestimation bias during training. The actor and critics are updated jointly to maximize the expected return while maintaining policy entropy, which promotes exploration during training. During training, the policy remains stochastic and samples actions from the Gaussian distribution. For closed-loop evaluation and real-time deployment, exploration is disabled and the actor is evaluated using its maximum-likelihood action. The deployed controller is therefore a deterministic actor-only neural network.

3.3 Reward model

The reward function used for training is defined as

$$R = \frac{1}{1 + \mathbf{W}^T\,\mathbf{J}}, \tag{20}$$

$$\mathbf{W} = \begin{bmatrix} w_s \\ w_{is} \\ w_a \end{bmatrix}, \qquad \mathbf{J} = \begin{bmatrix} \sum \ddot{x}^2 \\ \sum \left(\int \ddot{x}\,dt\right)^2 \\ \sum \delta_{\text{flap}}^2 \end{bmatrix}. \tag{21}$$

The first term penalizes the squared structural accelerations measured by the wing-mounted accelerometers. The second term penalizes the squared integrated accelerations and is included to discourage low-frequency flexible motion or drift. The acceleration integral is reset at the beginning of each training episode. The third term penalizes large flap deflections. The flap-deflection penalty is evaluated in degrees, while the acceleration terms are evaluated using SI units. No explicit flap-rate penalty is included in the reward function. This keeps the number of reward-tuning parameters low. Aggressive or high-frequency actuation is instead assessed during verification through actuator-dynamics and rate-limit checks.

The reward weights used for all controller variants are

$$w_s = 10^{-1}, \qquad w_{is} = 10^{-1}, \qquad w_a = \frac{10^{-1}}{n_a}, \tag{22}$$

where n_a is the number of active actuators. The scaling of w_a with n_a keeps the total actuator penalty comparable across different actuator configurations. The reciprocal form ensures that the reward remains bounded, with larger values corresponding to lower structural response and lower control effort.

3.4 Offline controller training

Table 2: Hyperparameters of the SAC agent.

Parameter	Value
Actor/Critic Optimizer	Adam
Actor Learning Rate α_a	1×10^{-4}
Critic Learning Rate α_c	1×10^{-4}
Discount Factor γ	0.9999
Target Smoothing Factor τ	1×10^{-3}
Mini-Batch Size	128
Experience Buffer Length	1×10^{6}
Warm-start Steps	1000
Sample Time T_s	0.001 s
Episode Duration	55 s
Training Episodes	500
Activation Functions	ReLU, Softplus (σ)

The SAC controller was trained offline using the linear ASE model at a freestream velocity of $30\ \mathrm{m/s}$. The Reinforcement Learning and Deep Learning Toolboxes from MathWorks [19] were used together with a customized `SACAgent` class. Further implementation details of the training and deployment workflow are reported in [20]. Each controller variant was trained for 500 episodes. Each episode lasted 55 s and used a linear frequency-sweep gust excitation from 0.1 Hz to 25 Hz over the first 50 s. The remaining 5 s contained zero gust excitation, allowing the controller to learn the desired zero-disturbance/zero-actuation behavior. Training was terminated after a fixed budget of 500 episodes. Parallel training was not used, since the complete training of one controller variant could be completed in less than three hours. Training was performed on a workstation equipped with an Intel Core i9-14900K processor, 128 GB RAM, and an NVIDIA GeForce RTX 4070 GPU, although no parallel training was used. The main hyperparameters used for training are summarized in Table 2.

As shown in Fig. 5, the episode reward and moving-average reward increase during the early part of training and eventually reach a plateau, suggesting convergence of the learned policy.

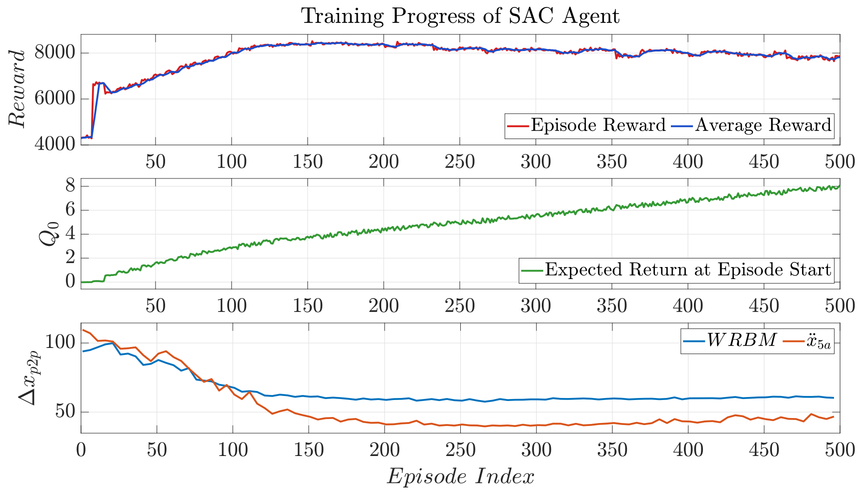


Figure 5: Training statistics of the SAC agent. Top: episode and moving-average rewards, showing convergence of the learning process. Middle: evolution of the expected return Q_0 at the beginning of each episode. Bottom: maximum peak-to-peak responses of WRBM and vertical acceleration $\ddot{x}_{5a}$ over each episode, showing progressive reduction of the structural response during training.

The critic-estimated expected return Q_0 also increases over the training process, which is consistent with the improvement in the learned value function. The maximum peak-to-peak values of the WRBM and the acceleration at sensor 5a decrease as training progresses, indicating that the learned policy reduces the dominant structural response in the training environment. Although the WRBM is not part of the reward function, it is monitored during training as an engineering performance metric.

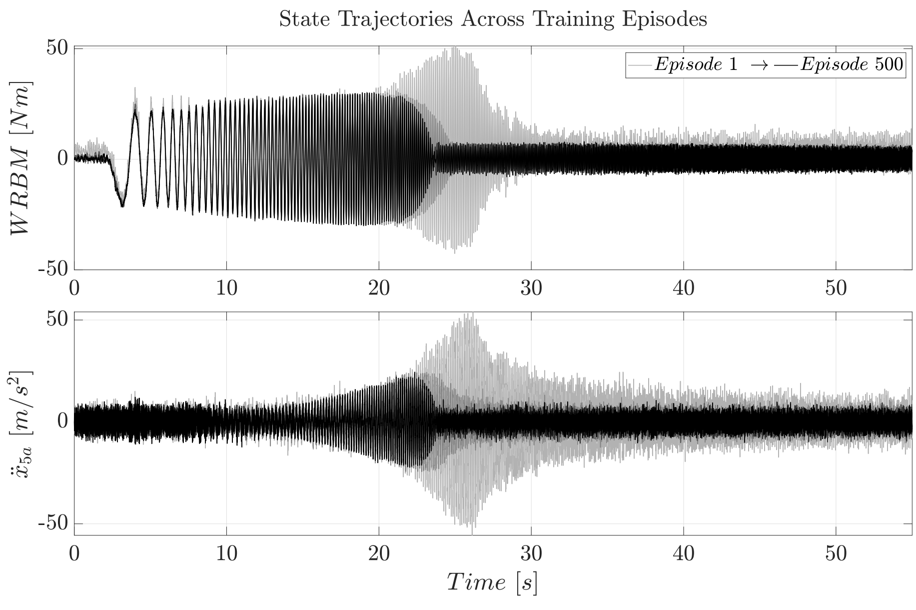


Figure 6: Evolution of state trajectories during training. Top: WRBM response over time, from early episodes shown in light gray to later episodes shown in darker traces. Bottom: vertical acceleration $\ddot{x}_{5a}$ at sensor 5a. Each episode is excited by a linear frequency-sweep gust from 0.1 Hz to 25 Hz over 50 s, followed by a zero-gust segment. The response reduction is particularly visible near the first bending mode around 8.6 Hz, which is excited approximately in the middle of the frequency sweep.

Figure 6 illustrates how the structural response changes over the course of training. In early episodes, the controller has not yet learned an effective policy and the gust excitation leads to large WRBM and acceleration amplitudes. As training progresses, the response amplitudes decrease, especially near the first wing bending mode. The later trajectories therefore show that the learned policy attenuates the dominant gust-induced aeroelastic response in the training environment. The trained actor was subsequently exported to Simulink and used in actor-only deterministic form for closed-loop evaluation. For the wind-tunnel experiments, the generated controller was deployed on the real-time control computer without additional controller-gain or actuator-gain tuning. The verification and validation process then introduced actuator dynamics, nonlinear actuator behavior, sensor filtering, rate and deflection limits, velocity variations, and finally real-time wind-tunnel testing. These steps are discussed in the following section.

4 SIMULATION-BASED VERIFICATION AND WIND-TUNNEL VALIDATION

This section presents the simulation-based verification and wind-tunnel validation results used to assess the experimentally deployed SAC-based GLA controllers. The broader verification and validation (V&V) framework, including the formulation of control-design specifications and the progressive implementation-verification workflow, is described in [21]. The present paper builds on that framework and focuses on the experimental assessment of controller performance on the SAFER^2 flexible-wing demonstrator.

4.1 Experimental setup

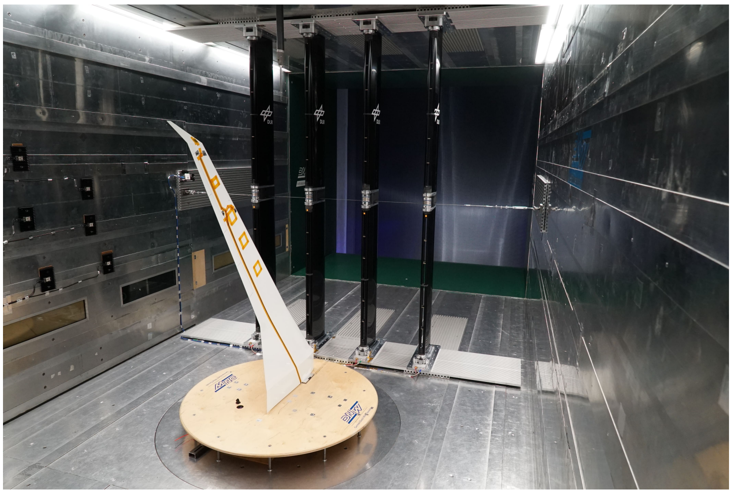


Figure 7: Experimental setup in the DNW–NWB low-speed wind tunnel showing the flexible aeroelastic wing demonstrator mounted on the turntable, together with the four rotating-slotted-cylinder gust-generator units and the ceiling-mounted five-hole probe used for inflow monitoring.

The experimental validation was conducted using the flexible aeroelastic wing demonstrator developed within the DLR project *SAFER^2* (Sensor and AI Fusion for Enhanced Performance and Reliability). The demonstrator originates from the earlier *oLAF* configuration and represents a scaled high-aspect-ratio transport-aircraft wing, with the corresponding reference aircraft having a span of 57.7 m and an aspect ratio of approximately 10 [22]. As shown in Fig. 7, the wind-tunnel model is installed as a cantilevered half-wing configuration inside the DNW–NWB test section. The wing is clamped at the root through a six-component piezoelectric balance, which provides measurements of the global aerodynamic forces and moments. In the present work, the WRBM obtained from this balance is used as the primary performance metric for assessing gust load alleviation.

The aeroelastic demonstrator is a modular composite wing equipped with four independently actuated trailing-edge flaps located on the outboard part of the span. These control surfaces enable different actuation configurations for active load alleviation and aeroelastic control studies.

The demonstrator also includes a comprehensive sensor suite for aeroelastic monitoring and sensor-model-fusion applications. The available instrumentation includes unsteady pressure sensors, acceleration sensors, a fiber-optical sensor system for strain measurements, Hall-effect sensors for local flap-deflection feedback, and the root balance for force and moment measurements [22]. For the SAC-based GLA experiments considered in this paper, the controller feedback is primarily based on four vertical acceleration sensors, where "vertical" denotes the aircraft-conventional z-axis, or out-of-plane, acceleration component of the wing. This distinction is important because the half-wing model is mounted upright in the wind tunnel. These sensors are located near the third and fourth flap stations, labelled 4a, 4b, 5a, and 5b, and capture the local bending and torsional response of the outboard wing section. In addition, measured flap positions from the Hall sensors are used in the feedback loop to account for the actual actuator response.

All tests were performed in the DNW–NWB low-speed wind tunnel in Braunschweig. The facility is a closed-loop wind tunnel of Göttingen type with a closed test section of $3.25\ \mathrm{m} \times 2.8\ \mathrm{m} \times 8.0\ \mathrm{m}$. The tunnel can provide freestream velocities up to $90\ \mathrm{m/s}$ at low turbulence levels [22, 23]. The model is mounted on the DNW turntable, which allows the model angle of attack to be adjusted. For all experiments presented in this paper, the angle of attack was fixed at $\alpha = 0°$. The main operating condition for the SAC-based GLA tests was a freestream velocity of $U_\infty = 30\ \mathrm{m/s}$ under atmospheric conditions.

Gust disturbances are generated by a dedicated gust-generator system installed upstream of the wing model. The system consists of four airfoil–rotating-slotted-cylinder units mounted between the tunnel floor and ceiling. Each unit contains a stationary airfoil followed by a rotating slotted cylinder driven by a brushless DC motor. By controlling the cylinder rotation, the system can generate repeatable continuous gust disturbances, which are approximately sinusoidal, as well as discrete gust events with a 1–cos like character [23]. The gust frequency is governed by the cylinder rotational speed, allowing excitation of the frequency range relevant to the flexible modes of the wing. Although four gust-generator units are available in the test setup, the experiments in this paper primarily use one- and two-generator excitation configurations. The specific gust-generator configuration used for each test case is introduced together with the corresponding results. The generated inflow conditions are monitored using an unsteady fast-response five-hole pressure probe mounted from the wind-tunnel ceiling.

The real-time control architecture consists of the wind-tunnel model, the measurement and data-acquisition systems, and the control computer. Sensor signals from the wing and the root balance are acquired and distributed to the data-acquisition and control systems. All relevant measurements are recorded using a DEWETRON data-acquisition system, while the trained SAC policy is executed on an ADwin Pro II real-time computer. The controller receives the selected acceleration and flap-position measurements, evaluates the neural-network policy in real time, and sends flap commands to the electromechanical actuators through PWM-based interfaces. Different actuator configurations are evaluated depending on the investigated SAC controller setup. This arrangement enables closed-loop experimental assessment of the trained SAC-based GLA controller under repeatable gust disturbances and allows direct comparison between open-loop and closed-loop wing-load responses.

4.2 Simulation-based verification

Figure 8 presents the simulated response of the single-actuator controller configuration, where only Flap 4 is used for gust load alleviation. The chirp gust is used as a broadband excita-

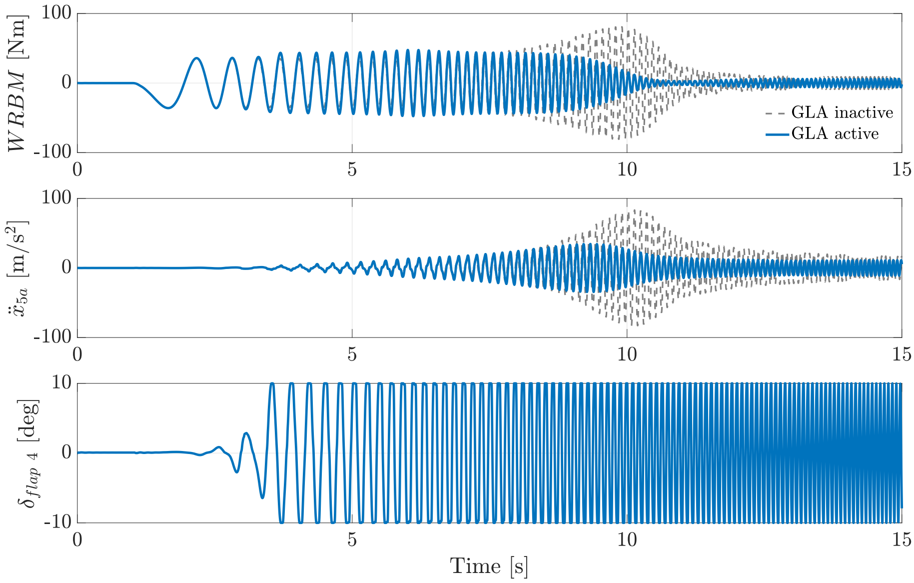


Figure 8: Time-domain simulation response for the single-actuator SAC-based GLA controller under chirp gust excitation.

tion to assess the controller response over a range of frequencies before wind-tunnel testing. In the open-loop case, denoted as GLA inactive, the WRBM and the acceleration response at sensor 5a increase as the chirp excitation approaches the dominant flexible dynamics of the wing. When the GLA controller is activated, both responses are visibly attenuated over the same time interval, indicating that the trained SAC policy reduces the gust-induced structural response in simulation. The corresponding Flap 4 deflection shows that the controller actively uses the available actuator authority during the most demanding part of the excitation. The deflection reaches the policy-imposed action bounds of approximately $\pm 10°$, which are introduced through the actor output scaling, although the physical actuator limits identified from ground tests are larger ($\pm 16°$, see Table 1). This indicates that the single-actuator configuration can provide effective attenuation, but also that the response is constrained by the action range selected during controller design.

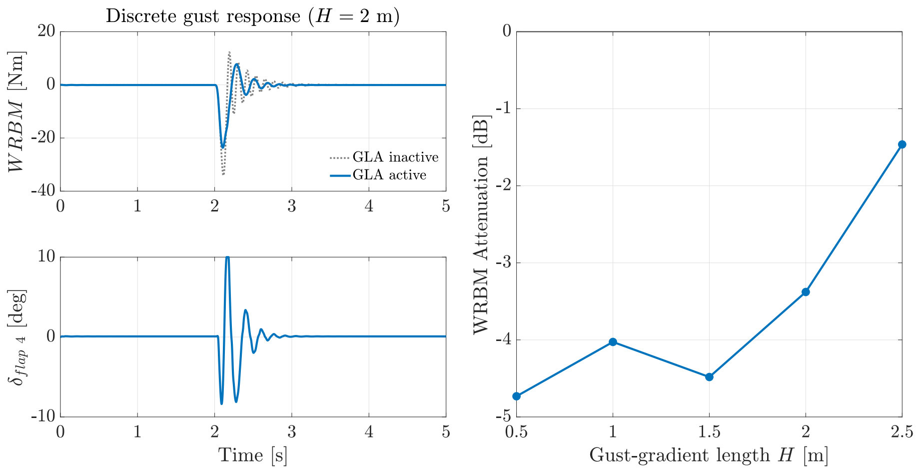


Figure 9: Simulation response of the single-actuator SAC-based GLA controller under discrete 1–cos gust excitation with varying gust-gradient length.

Figure 9 evaluates the single-actuator SAC-based GLA controller under discrete 1–cos gust excitation. For the representative case with $H = 2$ m, the controller reduces the peak-to-peak

WRBM response compared with the uncontrolled case and damps the subsequent transient oscillations. The corresponding Flap 4 deflection shows a short, high-authority control action around the gust encounter, with the response reaching the imposed deflection limits before rapidly decaying back toward the neutral position. The attenuation trend over different gust-gradient lengths indicates that the controller provides a consistent reduction in WRBM, since all attenuation values remain below $0\ \mathrm{dB}$. However, the reduction is modest and non-monotonic across the tested gust gradients. The strongest attenuation is obtained for shorter-to-intermediate gust gradients, whereas the benefit decreases for the longest case, $H = 2.5\ \mathrm{m}$. This suggests that the single-actuator policy provides more benefit for shorter-to-intermediate gust gradients than for slower gust encounters in the tested simulation cases.

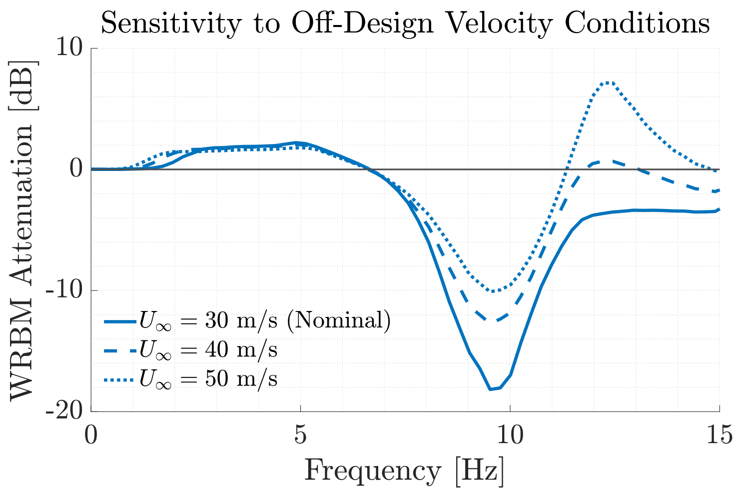


Figure 10: Empirical WRBM attenuation of the single-actuator SAC-based GLA controller under chirp excitation for nominal and off-design freestream velocities.

Figure 10 evaluates the sensitivity of the fixed-weight SAC policy to off-design freestream velocities. The attenuation is computed as the frequency-dependent ratio between the closed-loop and open-loop WRBM response, expressed in dB. At the nominal training condition of $U_\infty = 30\ \mathrm{m/s}$, the controller provides strong attenuation around the dominant flexible response region, with a maximum reduction of approximately 18–$19\ \mathrm{dB}$. A similar trend is retained at $40\ \mathrm{m/s}$, although the attenuation around the main response region is reduced. For both cases, only small amplification is observed at lower frequencies, remaining close to the $0\ \mathrm{dB}$ reference. At $50\ \mathrm{m/s}$, the controller still attenuates the response around the first dominant attenuation region; however, a clear amplification peak appears at higher frequencies, reaching approximately $7\ \mathrm{dB}$ near 12–$13\ \mathrm{Hz}$. This indicates a spillover region where the fixed policy no longer provides attenuation under the increased dynamic pressure. The result therefore identifies a practical operating limit of the nominally trained controller and motivates either velocity-conditioned policies or training over a broader freestream-velocity envelope.

Figure 11 compares four SAC-based GLA controllers trained with different actuator configurations. Such an analysis is important for GLA design because the achievable load reduction depends not only on the control law, but also on the spanwise location, number, and available authority of the control surfaces. The single-actuator case using Flap 4 already provides strong attenuation around the dominant flexible response region. However, adding inboard control surfaces changes both the attenuation bandwidth and the high-frequency behavior. The three- and four-actuator configurations achieve comparable peak attenuation near the main flexible mode, while reducing the high-frequency amplification observed for the one- and two-actuator

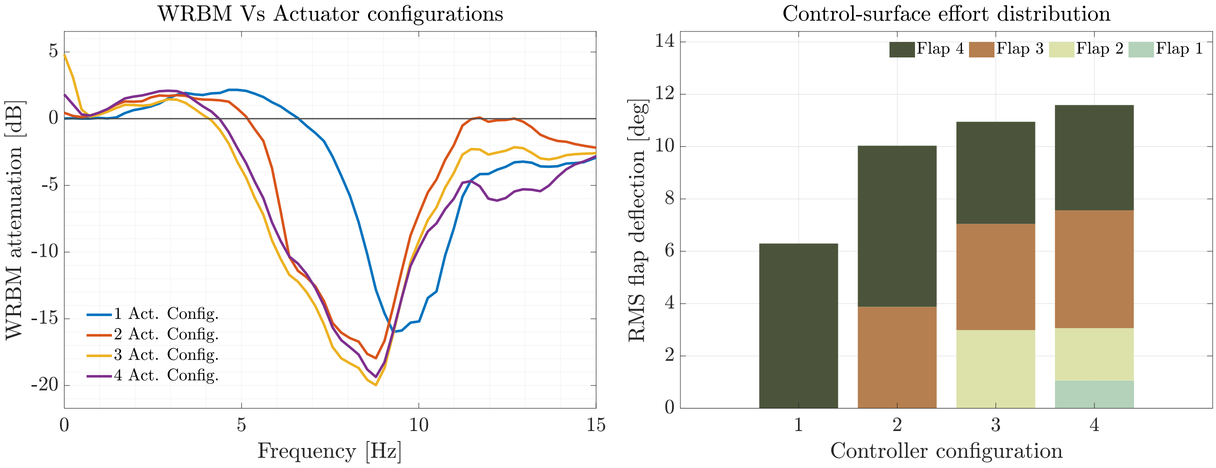


Figure 11: Frequency-domain WRBM attenuation and RMS flap-deflection effort for four SAC-based GLA controller configurations with increasing actuator availability.

cases above approximately 14 Hz. The control-effort distribution shows that the outer flaps remain dominant, especially Flap 4 and Flap 3, whereas Flap 1 contributes only weakly in the four-actuator configuration. This indicates that the outboard surfaces are most effective for exciting and counteracting the relevant bending-dominated load response. From a control-design perspective, the result illustrates a useful property of the RL-based formulation for distributed GLA. Once the training environment, reward structure, and sensor set are defined, different actuator configurations can be investigated mainly by changing the action dimension and retraining the policy. This provides a scalable workflow for exploring control-surface combinations with limited control-structure redesign effort. At the same time, the comparison shows that simply adding more actuators does not guarantee proportional improvement, and that actuator selection remains an important design variable.

4.3 Wind-tunnel validation

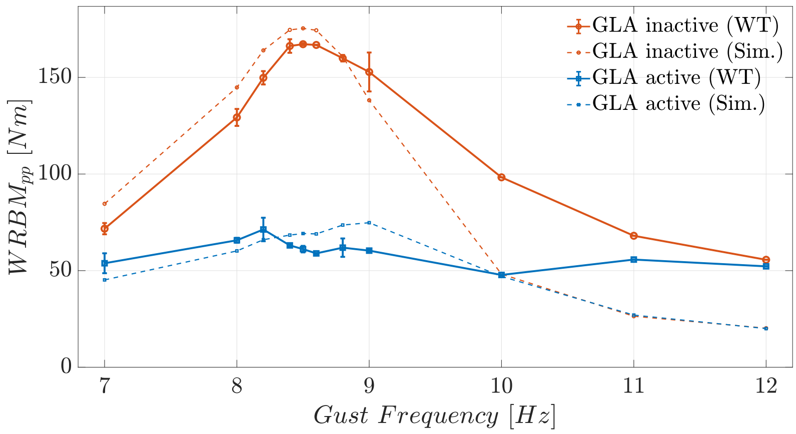


Figure 12: Wind-tunnel validation of the single-actuator SAC-based GLA controller under harmonic gust excitation using one gust generator. Error bars indicate the standard deviation across repeated wind-tunnel runs.

Figure 12 presents the wind-tunnel validation results for the single-actuator controller configuration, where Flap 4 is used for gust load alleviation. Each test point corresponds to a pure sinusoidal gust disturbance of fixed amplitude at a different excitation frequency. The wind-tunnel data are reported as mean peak-to-peak WRBM values across repeated runs, with the error bars denoting the corresponding standard deviation. Two repeated runs were performed for most frequencies, while three runs were available around the first bending-mode region. The uncontrolled response exhibits a clear resonance around 8.4–8.8 Hz, where the measured peak-to-peak WRBM reaches approximately 165–170 Nm. Activating the SAC-based GLA controller reduces the response to about 60 Nm in the same frequency range. At 8.6 Hz, this corresponds to an experimentally measured load reduction of approximately 65%. The simulation results, which include the nonlinear actuator and sensor models introduced in Table 1, reproduce the main resonance location and the closed-loop attenuation trend. This agreement supports the usefulness of the implementation-aware simulation model for pre-test verification. At the same time, noticeable differences remain between simulation and experiment, particularly away from the main resonance region. Since these differences appear in both the GLA-inactive and GLA-active cases, they are likely dominated by residual plant and gust-field modeling uncertainties rather than by the controller alone. Experimental effects, including measurement noise and the sensitivity of peak-to-peak processing at lower response amplitudes, may also contribute, especially at the higher tested frequencies.

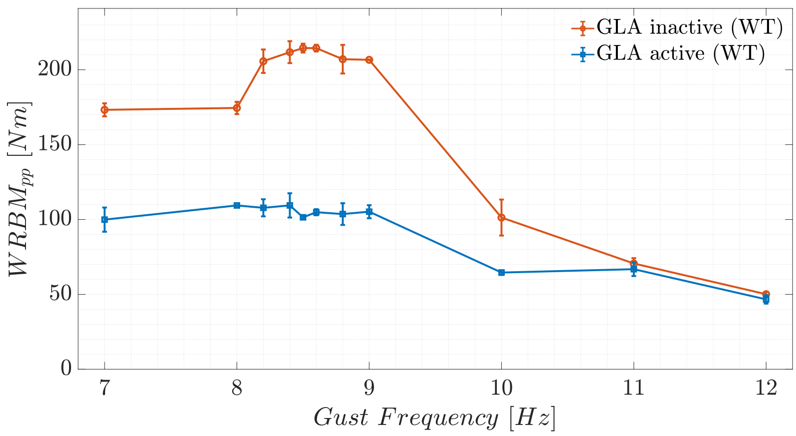


Figure 13: Wind-tunnel validation of the single-actuator SAC-based GLA controller under increased harmonic gust intensity using two gust generators. Error bars indicate the standard deviation across repeated wind-tunnel runs.

Figure 13 evaluates whether the same trained single-actuator SAC policy preserves its load-alleviation capability under increased external gust forcing. In contrast to the one-generator case in Fig. 12, this test uses two gust generators while keeping the controller, feedback signals, and action limits unchanged. The uncontrolled response again exhibits a dominant peak around the first bending-mode region, with the peak-to-peak WRBM increasing to approximately 214 Nm at 8.6 Hz. With the GLA controller active, the corresponding response is reduced to approximately 105 Nm, giving an experimentally measured load reduction of about 51%. The result shows that the controller retains substantial attenuation under stronger harmonic gust excitation, but with a lower relative reduction than in the one-generator case. This behavior is consistent with the finite authority of the single-actuator configuration: as the disturbance intensity increases, the controller can still reduce the resonant load response, but the achievable attenuation becomes limited by the available control action. The small standard deviations across repeated runs indicate good experimental repeatability, particularly near the critical frequency range where three repeated measurements were available. Therefore, this test complements the nominal one-generator validation by providing experimental evidence of retained attenuation under increased gust intensity and good repeatability.

Figure 14 presents a wind-tunnel comparison of the four actuator configurations previously

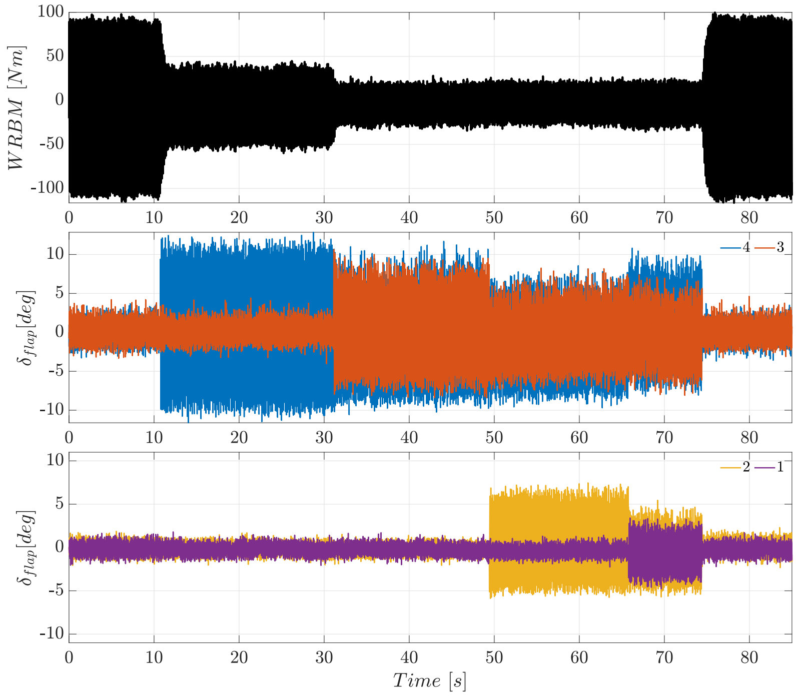


Figure 14: Wind-tunnel comparison of four separately trained SAC-based GLA controller configurations under constant harmonic gust excitation at 8.6 Hz using two gust generators. The configurations are activated sequentially from Flap 4 only to Flaps 1–4.

studied in simulation in Fig. 11. The test was conducted under a constant harmonic gust excitation at 8.6 Hz using two gust generators, corresponding to the critical first bending-mode region. Four separately trained SAC controllers were activated sequentially during the same run: first using Flap 4 only, followed by Flaps 3–4, Flaps 2–4, and finally Flaps 1–4. The beginning and final parts of the time history correspond to the GLA-inactive condition; the small non-zero flap signals in these intervals are attributed to measurement noise. The sharp reduction in WRBM after controller activation confirms that the trained policies can be deployed on the wind-tunnel hardware and provide immediate load attenuation under the same gust condition. The result also supports the actuator-configuration trends observed in simulation. The largest change occurs when the outboard Flap 4 controller is activated, while adding further inboard control surfaces does not produce a proportional additional reduction in WRBM. The measured flap histories show that the outer flaps remain the dominant contributors, whereas the inboard surfaces are used with lower amplitudes and mainly redistribute control effort. This is consistent with the simulation-based observation that the outboard surfaces are most effective for controlling the bending-dominated load response. The smooth sequential activation is additionally useful from an experimental GLA perspective, because it demonstrates that different trained policies can be transitioned on the real-time system without visible transient growth or loss of attenuation.

4.4 Additional results

Figure 15 shows an additional wind-tunnel test used to assess the influence of actuator rate capability on GLA performance. The test was conducted with the single-actuator controller using Flap 4, under harmonic gust excitation at 8.6 Hz with one gust generator. The initial and final parts of the time history correspond to the GLA-inactive condition, where the WRBM response is high. After controller activation, the response is strongly attenuated. At $t = 21$ s, an artificial rate limitation is introduced at the actuator command interface, reducing the allowable

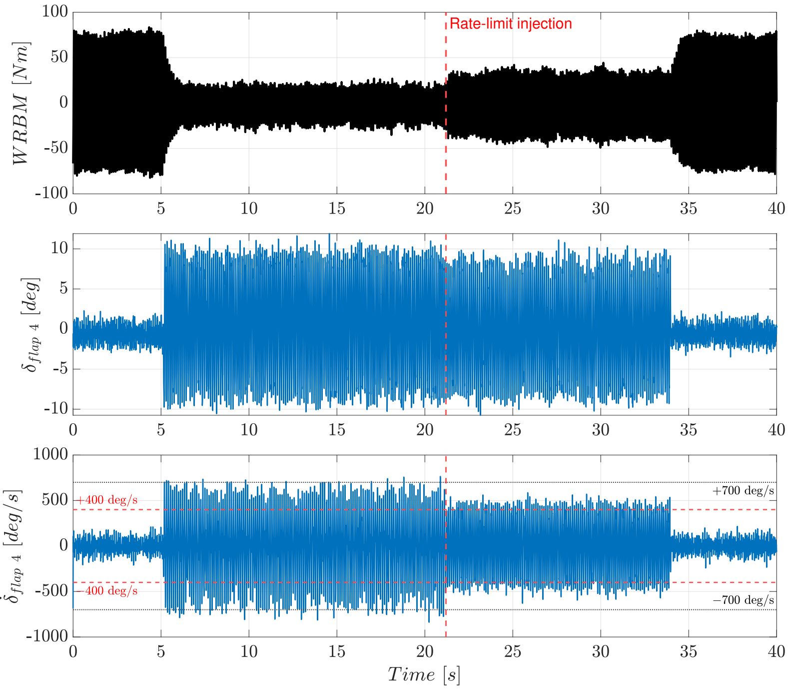


Figure 15: Effect of artificial actuator-rate limitation on the wind-tunnel performance of the single-actuator SAC-based GLA controller. The Flap 4 rate limit is reduced at the command interface to 400 deg/s during harmonic gust excitation at 8.6 Hz using one gust generator.

Flap 4 rate to 400 deg/s, which is below the identified physical actuator rate capability of approximately 700 deg/s from Table 1. After the rate-limit injection, the measured WRBM amplitude increases, indicating a degradation in load-alleviation performance due to reduced actuator rate authority. The flap-deflection amplitude remains within the same general range, but the estimated flap-rate signal shows that the actuator can no longer follow the faster control action available before the injection. This is important for GLA, where the phase and timing of the control-surface motion are critical near the dominant aeroelastic frequency. Nevertheless, the response remains substantially below the GLA-inactive level, showing that the closed-loop response remains bounded in this test and continues to provide partial attenuation despite the imposed rate constraint.

Overall, the simulation and wind-tunnel results show that the fixed SAC policy can attenuate the dominant gust-induced WRBM response on the experimental hardware. The comparison between simulation and experiment supports the use of implementation-aware simulation for pre-test verification, while the wind-tunnel tests show repeatable attenuation under harmonic gust excitation. The results also indicate that the achievable load reduction depends on gust intensity and actuator authority, and that implementation effects such as actuator-rate limits must be considered before deployment. Furthermore, under comparable experimental conditions using a dual-actuator setup, a robust control design utilizing μ-synthesis achieved a relative WRBM reduction of approximately 66% at the first bending mode frequency [24]. While the SAC agent provides higher peak attenuation of 80%, it was trained and verified only up to 25 Hz. In contrast, the μ-synthesis controller explicitly accounts for broader bandwidth within the robust stability framework, while simultaneously enforcing low-frequency separation from the rigid-body control loop. This suggests a performance-robustness tradeoff, where the SAC method prioritizes peak performance within a narrower targeted bandwidth.

5 CONCLUSIONS AND FUTURE WORK

This paper demonstrated the design, offline training, simulation-based verification, and wind-tunnel deployment of a fixed deterministic SAC-based GLA controller for a flexible aeroelastic wing demonstrator. The results show that a neural-network policy trained on a linear ASE model can be transferred to the wind-tunnel setup and can provide substantial WRBM attenuation under harmonic gust excitation. For the single-actuator Flap 4 configuration, the controller achieved approximately 65% WRBM reduction at the first bending-mode frequency. Under increased gust excitation using two generators, the same fixed policy retained a 51% reduction. The comparison between simulation and experiment further indicates that experimentally identified actuator and sensor effects are important for implementation-aware pre-test assessment before wind-tunnel testing. At the same time, the results highlight practical aspects that require further investigation, including sensitivity to gust intensity, actuator authority, actuator-rate limitations, off-design operating conditions, and the distribution of control effort across multiple trailing-edge flaps.

Future work will extend this framework in several directions. A natural next step is the transition from wind-tunnel testing to in-flight validation, for example on a small uncrewed aircraft or an experimental research aircraft, where the interaction between rigid-body flight control and aeroelastic load alleviation must be addressed. Further work will also investigate training with richer data sources, including nonlinear ASE models, computational-fluid-dynamics or fluid–structure-interaction simulations, and plant-derived wind-tunnel or flight-test data. In particular, offline retraining or fine-tuning with experimental data is a promising route to improve robustness and generalization while keeping the deployed controller fixed during operation and therefore more amenable to verification. The same learning-based framework could also be extended to AFS, where model dependence, nonlinear aeroelastic behavior, and sensor/actuator selection pose similar challenges. Finally, alternative sensing and actuation concepts, such as lidar-based gust preview, pressure sensors, fiber Bragg grating strain measurements, piezoelectric sensors, and morphing or distributed control surfaces, should be explored to assess how RL can exploit complementary information sources for future active aeroelastic control systems.

ACKNOWLEDGMENTS

This work was conducted within the Sensor and AI Fusion for Enhanced Performance and Reliability (SAFER^2) project of the German Aerospace Center (DLR). The authors gratefully acknowledge all contributors involved in the planning, preparation, and execution of the wind-tunnel experiment, which made the integrated control tests possible. Special thanks are extended to Manuel Pusch, Tijmen Pollack, João Paulo Dias Martins, Anna Altkuckatz, Marc Braune, Holger Mai, Johannes Dillinger, Robin Volkmar, Keith Soal, Martin Tang, Marc Böswald, Simon Schulz, Thiemo Kier, and the DNW–NWB wind-tunnel team.

6 REFERENCES

[1] Wright, J. R. and Cooper, J. E. (2014). *Introduction to Aircraft Aeroelasticity and Loads*. Chichester, UK: John Wiley & Sons, 2 ed. doi:10.1002/9781118700440.

[2] Hoblit, F. M. (1988). *Gust Loads on Aircraft: Concepts and Applications*. Washington, DC, USA: American Institute of Aeronautics and Astronautics. doi:10.2514/4.861888.

[3] Regan, C. D. and Jutte, C. V. (2012). Survey of applications of active control technology for gust alleviation and new challenges for lighter-weight aircraft. Tech. Rep. NASA/TM-2012-216008, NASA.

[4] Federal Aviation Administration (2014). Advisory circular ac 25.341-1: Dynamic gust loads. Tech. rep., U.S. Department of Transportation, Washington, DC, USA.

[5] European Union Aviation Safety Agency (2023). Certification specification and acceptable means of compliance for large aeroplanes (cs-25), amendment 28.

[6] Pusch, M., Ossmann, D., Dillinger, J., et al. (2019). Aeroelastic modeling and control of an experimental flexible wing. In *AIAA Scitech 2019 Forum*. San Diego, CA, USA. doi:10.2514/6.2019-0131.

[7] Stalla, F., Kier, T. M., Looye, G., et al. (2024). Wind tunnel testing of active gust load alleviation of a flexible wing. In *Proceedings of the International Forum on Aeroelasticity and Structural Dynamics (IFASD 2024)*. The Hague, Netherlands.

[8] Pusch, M., Konatala, R., Wüstenhagen, M., et al. (2026). Flight testing blending-based active flutter suppression. In *Proceedings of the 23rd IFAC World Congress*. To appear.

[9] Cavaliere, D. and Fezans, N. (2024). A practical approach to automated multiobjective gust load alleviation control design in a structured h2/h-infinity framework. In *CEAS EuroGNC 2024*. Bristol, UK. doi:10.82124/CEAS-GNC-2024-042.

[10] Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. doi:10.1016/0893-6080(89)90020-8.

[11] Vidyasagar, M. (2003). *Learning and Generalisation: With Applications to Neural Networks*. London: Springer, 2 ed. doi:10.1007/978-1-4471-3748-1.

[12] Haarnoja, T., Zhou, A., Abbeel, P., et al. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. doi:10.48550/arXiv.1801.01290.

[13] Haughn, K. P. T., Harvey, C., and Inman, D. J. (2024). Deep learning reduces sensor requirements for gust rejection on a small uncrewed aerial vehicle morphing wing. *Communications Engineering*, 3(1), 136. doi:10.1038/s44172-024-00201-8.

[14] Hu, Y., Dai, Y., Li, J., et al. (2025). Reinforcement learning for gust load control of an elastic wing via camber morphing at arbitrary sinusoidal gusts. *Aerospace Science and Technology*, 110174. doi:10.1016/j.ast.2025.110174.

[15] Kier, T. M. and Hofstee, J. (2004). VarLoads – eine simulationsumgebung zur lastenberechnung eines voll flexiblen, freifliegenden flugzeugs. In *Deutscher Luft- und Raumfahrtkongress (DLRK)*. Dresden, Germany: DGLR.

[16] Kier, T. and Looye, G. (2009). Unifying manoeuvre and gust loads analysis models. In *IFASD*. Seattle, USA.

[17] Stalla, F., Kier, T. M., Looye, G., et al. (2024). Aeroservoelastic modeling and robust control for gust load alleviation of an experimental wing. In *AIAA SCITECH 2024 Forum*. doi:10.2514/6.2024-1442.

[18] Tang, M., Böswald, M., Govers, Y., et al. (2021). Identification and assessment of a nonlinear dynamic actuator model for controlling an experimental flexible wing. *CEAS Aeronautical Journal*, 12(2), 413–426. doi:10.1007/s13272-021-00504-y.

[19] MathWorks (2024). Matlab reinforcement learning toolbox. Version R2024b.

[20] Konatala, R., Looye, G., and Kampen, E.-J. V. (2026). Soft actor-critic based gust load alleviation control design for an experimental flexible wing. In *Proceedings of the European Control Conference (ECC)*. Reykjavik, Iceland, p. To appear.

[21] Konatala, R., Stalla, F., Kier, T., et al. (2026). Verification and validation of reinforcement learning-based aeroelastic control system. In *Proceedings of the 23rd IFAC World Congress*. Busan, South Korea, p. To appear.

[22] Schmidt, T. G., Volkmar, R., Micheli, B., et al. (2025). Assessing sensor-model-fusion technologies on a flexible aeroelastic wing demonstrator through wind tunnel testing in the dlr project safer². In *DLRK 2025*.

[23] Schmidt, T. G., Dillinger, J., Ritter, M. R., et al. (2024). Design and experimental characterization of a gust-generator concept with rotating-slotted cylinders in the low-speed wind tunnel dnw-nwb. In *IFASD 2024*.

[24] Stalla, F., Looye, G., Pusch, M., et al. (2026). Integrating aeroelastic and primary flight control: Robust design and wind tunnel demonstration. In *AIAA SCITECH 2026 Forum*. doi:10.2514/6.2026-1557.