ROTOR BLADE DYNAMIC STALL: FROM LARGE EDDY SIMULATIONS TO REDUCED ORDER MODELS

Giacomo Baldan
Politecnico di Milano
Milan, Italy

Francesco Manara
Helicopter Division, Leonardo SpA
Samarate, Italy

Gregorio Frassoldati
Università Degli Studi Roma Tre
Rome, Italy

Carlo Cassinelli
Helicopter Division, Leonardo SpA
Samarate, Italy

Alberto Guardone
Politecnico di Milano
Milan, Italy

ABSTRACT

We investigate deep dynamic stall over a pitching NACA0012 airfoil using wall-resolved Large Eddy Simulations at Reynolds number 135 000 and reduced frequency 0.1. The influence of the span-to-chord ratio on unsteady aerodynamic behavior is assessed through simulations with varying span lengths from $0.2c$ to $1.2c$. Results demonstrate that smaller spans introduce significant two-dimensional effects, affecting lift, drag, and vortex dynamics. The simulation data are compared against experimental measurements and state-of-the-art RANS and hybrid RANS/LES methods. Furthermore, a reduced-order model based on deep learning is developed using a flow-matching diffusion approach with a DiT architecture. This model successfully predicts flow fields across a wide range of operating conditions, with high fidelity and a fraction of the computational cost. The integration of high-fidelity simulations and efficient ROMs offers a robust pathway toward real-time aeroelastic simulations and advanced control strategies in rotorcraft design.

INTRODUCTION

Dynamic stall is a transient, unsteady aerodynamic phenomenon that occurs when airfoils experience rapid changes in their angle of attack. Unlike static stall, where lift suddenly decreases once the airfoil exceeds a critical stall angle, dynamic stall involves a sequence of complex events. Initially, the rapid angle change triggers the separation of the attached boundary layer, leading to the formation of a dynamic stall vortex (DSV) that convects along the suction surface. As this vortex develops, it interacts intricately with the boundary layer. At angles exceeding the static stall limit, the dynamic stall mechanism induces extensive flow separation over a large portion of the airfoil. This flow separation initially results in a transient increase in lift over the static limit, which, at high forward speeds, can contribute to helicopter trim. However, this is invariably followed by a pronounced and often delayed loss of lift, accompanied by a substantial increase in aerodynamic drag and a significant nose-down pitching moment. These unsteady aerodynamic loads can severely affect the performance and structural stability of the system. Consequently, dynamic stall is a critical factor in the design and operation of various aerodynamic structures, notably in helicopter rotor blades, wind turbines, and during high-performance aircraft maneuvers. In rotorcraft, the unsteady aerodynamic loads produced by dynamic stall can lead to significant structural fatigue, increased vibration, and elevated noise levels, ultimately affecting both performance and the operational lifespan of the rotor system. Moreover, dynamic stall remains one of the major limiting factors in achieving maximum advancing speeds (Ref. 1).

The inherently complex and transient nature of dynamic stall poses a substantial challenge in the modeling, simulation, and design of aerodynamic systems operating under such conditions (Ref. 2). To address this, research efforts are largely directed toward advancing high-fidelity Computational Fluid Dynamics (CFD) tools and developing low-fidelity reduced-order models (ROMs) for effective flow control strategies. A deeper understanding of dynamic stall and its effects is crucial for refining the design and control of aerodynamic structures. Ultimately, these advancements contribute to safer and more efficient flight, enhanced wind energy generation, and improved aerodynamic performance across a wide range of operating conditions.

A substantial body of research on dynamic stall employs Unsteady Reynolds-Averaged Navier-Stokes (URANS) models and hybrid RANS/LES techniques (Ref. 3). These studies generally find that discrepancies between experimental results and numerical simulations often stem from the turbulence model used to close the RANS equations. A wide range of turbulence models has been analyzed, from the one-equation Spalart-Allmaras model and full Reynolds stress models (RSM) to the widely adopted $k-\varepsilon$ and $k-\omega$ SST models. Additionally, transition models, such as the recent $k-\omega$ SST with the intermittency equation, are also in use.

In the domain of hybrid strategies, detached eddy simulations (DES) along with its improved versions (DDES, IDDES) are common, and more recent methods like scale adaptive simulations (SAS) and stress-blended eddy simulations (SBES) have been introduced.

Large Eddy Simulations (LES) are also extensively used to investigate dynamic stall. Research using LES has explored various parameters for finite wings, including aspect ratio, sweep angle, and compressibility effects, as well as studies on nominal infinite wings via periodic spanwise boundary conditions. Best practices in DES, as outlined in (Ref. 4), recommend a spanwise extent of at least one chord length to accurately capture the DSV, whose size is experimentally known to be roughly equivalent to the airfoil chord during deep dynamic stall (Ref. 5). Despite these guidelines, many LES studies still use extruded airfoil geometries with spanwise lengths between only 5% and 20% of the chord length (Ref. 6).

Given the high computational expense of dynamic stall CFD simulations, ROMs are essential for reducing the cost of analyses such as aeroelastic evaluations and flight mechanics studies. Currently, the Leishman-Beddoes and ONERA models are the most widely used. The Beddoes model originated in the 1970s with an indicial formulation. The latest version, known as the Third Generation Model (Ref. 7), refines load prediction accuracy at low Mach numbers and reduced frequencies by introducing two additional equations compared to earlier versions. In contrast, the ONERA model describes unsteady airfoil behavior through a set of nonlinear differential equations that predict time-domain responses. Rather than employing physics-based modeling, the ONERA approach emphasizes capturing the hysteresis characteristics of dynamic stall. The most recent version, referred to as the ONERA BH model, is presented in reference (Ref. 8).

This work employs a wall-resolved LES framework to simulate a NACA0012 airfoil undergoing a complete sinusoidal pitching cycle at Reynolds number Re = 135 000 and reduced frequency $k = \omega c/2V_\infty = 0.1$, where ω is the pitching frequency. The primary objective is to quantify the influence of the spanwise dimension on the aerodynamic response during both the upstroke and downstroke phases. Span lengths ranging from $0.2c$ to $1.2c$ are examined to assess how dynamic stall characteristics evolve with span variation. The simulation results are compared with experimental data from Lee et al. (Ref. 9) and against state-of-the-art RANS and hybrid RANS/LES models, in order to determine whether the discrepancies observed arise from the turbulence model employed or from other factors, such as surface roughness or wind tunnel effects. Additionally, this study introduces a reduced-order model based on deep neural networks to predict force and moment coefficients throughout the entire pitching cycle, encompassing small oscillations as well as both light and deep dynamic stall conditions. By employing diffusion models, it captures not only the average profile behavior but also the stochastic cycle-to-cycle variations. The primary objective is to develop a model that efficiently computes aerodynamic loads for integration into an aeroelastic simulation framework. Unlike semiempirical models that rely on predefined laws, such as the Leishman-Beddoes one, and require tuning for each operating condition, the present approach assumes no constraints on the flight conditions.

NUMERICAL SIMULATION SETUP

The present numerical simulations are designed to replicate the experimental conditions described by (Ref. 9). Specifically, the study focuses on a NACA0012 airfoil subjected to a sinusoidal pitching motion, defined by:

$$\alpha(t) = 10^\circ + 15^\circ \sin\left(18.67\,\mathrm{Hz} \cdot t\right) \tag{1}$$

The simulations are performed at a Reynolds number of $\mathrm{Re} = 1.35 \times 10^5$ and a reduced frequency of $k = \omega c/2V_\infty = 0.1$. The free-stream velocity is fixed at $V_\infty = 14\,\mathrm{m/s}$, with the ambient pressure maintained at $P_\infty = 1\,\mathrm{atm}$.

The computational mesh employed in this study is generated using *Cadence® Fidelity Pointwise v2023.2*. To avoid non-physical flow features often introduced by strongly anisotropic elements in conventional C-grid topologies, particularly near the trailing edge (Ref. 3), an O-grid configuration is adopted. Table 1 summarizes the properties of the two-dimensional mesh. The curvilinear coordinate ξ develops along the airfoil surface, while η denotes the wall-normal direction. The O-grid is constructed via a hyperbolic extrusion process originating from the airfoil surface, with a geometric growth rate of 1.03 governing the wall-normal spacing. The mesh satisfies established resolution guidelines for wall-resolved large-eddy simulations (Refs. 10, 11), which typically recommend $\xi^+ < 40$, $\eta^+ < 5$, and $z^+ < 20$ near the wall. These criteria are met across all mesh configurations, even under the most demanding flow conditions, such as high angles of attack where local velocities exceed twice the free-stream value. Accordingly, the first grid point is placed at a distance of $10^{-4}\,c$ from the wall. Special attention is given to element sizing near the leading and trailing edges, regions critical for resolving dynamic stall. The leading edge is particularly important due to the presence of a laminar separation bubble during pitch-up, where the DSV originates. The trailing edge is equally crucial, as it often marks the beginning of flow separation. To improve mesh quality and convergence behavior, the blunt trailing edge is connected to a circular arc, enhancing element orthogonality and avoiding the imposition of an artificial separation point. At the airfoil nose, the streamwise grid spacing is $\Delta\xi_{\text{leading edge}} = 6.67 \times 10^{-4}\,c$, while at the trailing edge it is $\Delta\xi_{\text{trailing edge}} = 4.35 \times 10^{-4}\,c$. To reduce the mesh size, a coarser resolution is used on the pressure side, which plays a less fundamental role in DSV advection. The suction side is discretized with 550 nodes, yielding an average spacing of $\overline{\Delta\xi}_{\text{suction}} = 1.88 \times 10^{-3}\,c$, whereas the pressure side uses 385 nodes with $\overline{\Delta\xi}_{\text{pressure}} = 2.69 \times 10^{-3}\,c$. All meshes extend to a farfield boundary located at $50c$ from the airfoil to minimize domain-induced interference. The three-dimensional grid is generated by uniform extrusion of the two-dimensional O-grid in the spanwise (z) direction. The number of nodes in this direction is selected to meet LES resolution criteria (Ref. 12). For the smallest spanwise extent of $0.2c$,

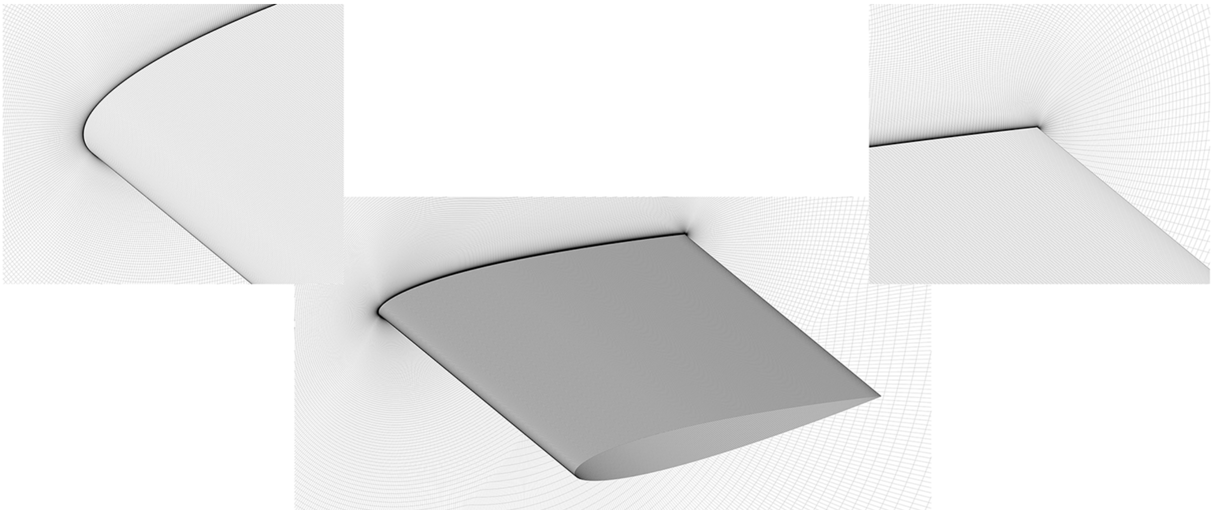

Figure 1. Details of the three-dimensional mesh of the wing section with span-to chord ratio of 1.2.

Table 1. 2D O-grid specifications. The size of the elements is made dimensionless by the chord size *c*.

N_ξ	N_η	Nodes	Quads	$\Delta\xi_{LE}$	$\Delta\xi_{TE}$	$\overline{\Delta\xi}_{\text{suction}}$	$\overline{\Delta\xi}_{\text{pressure}}$
934	334	311 956	311 622	$6.67 \cdot 10^{-4}$	$4.35 \cdot 10^{-4}$	$1.88 \cdot 10^{-3}$	$2.69 \cdot 10^{-3}$

Table 2. Adopted grid specifications. Resolution in z direction is kept constant.

Span size	N_z	Nodes	Hexahedra
$0.2c$	201	62.7M	62.1M
$0.4c$	401	125.1M	124.3M
$0.8c$	801	249.9M	248.6M
$1.2c$	1201	374.7M	372.8M

201 points are used. Maintaining constant spanwise resolution, the grids for extents of $0.4c$, $0.8c$, and $1.2c$ employ 401, 801, and 1201 nodes, respectively. Details of these configurations are listed in Table 2.

The incompressible LES equations are solved using *ANSYS Fluent 2023R2*. Spatial discretization employs a second-order upwind scheme, with gradient reconstruction via a least-squares cell-based method. Fluxes are computed using the Rhie-Chow momentum interpolation. Time integration is performed using a bounded second-order implicit scheme, and pressure-velocity coupling is handled through the SIMPLEC algorithm. The WALE subgrid-scale model is employed to ensure appropriate near-wall eddy viscosity behavior without requiring empirical damping functions. The pitching motion is discretized over 144,000 time steps, maintaining a CFL number near unity. The airfoil undergoes rigid-body rotation, with the angular velocity prescribed by a user-defined function:

$$\dot{\alpha}(t) = \omega\, \alpha_s \cos(\omega t),$$

where α_s is the pitch amplitude. To apply the mean angle of attack α_0, the mesh is rigidly rotated before the unsteady simulation begins. To accelerate convergence, the simulation proceeds in three stages. A steady RANS solution at α_0 is first computed as a precursor step to initialize the flow field. Next, the unsteady pitching simulation is initiated with a negative pitch rate, and the initial quarter-cycle of pitch-down motion is discarded (Ref. 6). Finally, a complete pitch cycle is simulated, and the resulting flow field is analyzed.

Table 3. Range of conditioning inputs defining the operating conditions.

Variable	Min	Max
Mach	0.3	0.5
α_0	5°	10°
α_s	5°	10°
k	0.05	0.1

Neural Network Dataset Generation

The datasets used for training and testing the neural network are generated by solving the unsteady, compressible, two-dimensional RANS equations around a sinusoidally pitching NACA0012 airfoil. The numerical setup follows best practices established in prior studies (Refs. 13,14) and follows the best practice outlined for the LES simulation.

The design space is defined as a four-dimensional hypercube, where each dimension corresponds to a conditioning input for the neural network. Free-stream conditions are characterized by the Mach number, while the Reynolds-to-Mach ratio remains constant due to a fixed airfoil chord and a uniform free-stream temperature of 15°C. The remaining inputs describe the airfoil's motion: the mean angle of attack (α_0), pitching amplitude (α_s), and reduced frequency ($k = \omega c / 2V_\infty$). The ranges of these parameters are listed in Table 3. To ensure comprehensive and uniform coverage of the design space, we employ Halton's low-discrepancy sequence to sample operating conditions. Unlike purely random sampling, low-discrepancy sequences offer superior space-filling properties by minimizing clustering, particularly in high-dimensional domains (Ref. 15). The hypercube is sampled at 128 points for training and 16 points for testing, with different random seeds used to avoid overlap between datasets. Each point represents a nominal operating condition. To introduce variability, we generate 32 perturbed instances per nominal condition using multiplicative Gaussian noise:

$$x_{\text{perturbed}} = (1 + \mathcal{N}(0, 0.02))\, x_{\text{nominal}},$$

where $\mathcal{N}(0, 0.02)$ denotes a normal distribution with zero mean and a standard deviation of 0.02. This results in 4 096 simulations for training and 512 simulations for testing. Each

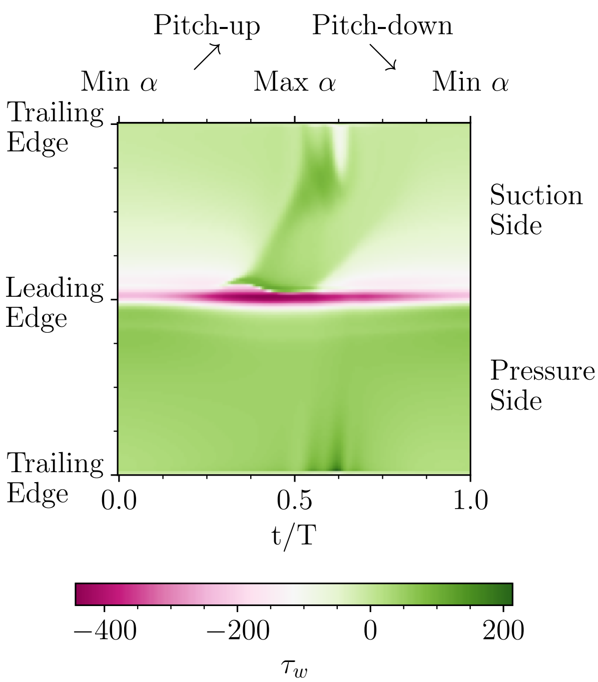

Figure 2. Spatio-temporal contour of the post-processed skin friction coefficient (τ_w) over one complete pitching cycle.

simulation output is stored as a three-dimensional array with dimensions corresponding to time steps, discrete surface locations, and physical variables. For computational efficiency, each field is downsampled to a resolution of 128×128. The saved variables include absolute pressure, density, temperature, signed skin friction, and tangential velocity gradients along the airfoil surface. This approach is mesh-independent, relying solely on surface data. Skin friction values are post-processed to reflect directional information. The sign is determined by the orientation of the τ_w vector: positive if aligned with the counterclockwise curvilinear coordinate (from trailing edge, around the airfoil, back to the trailing edge), and negative otherwise. This signed convention allows in identifying regions of flow separation and stall, providing informative inputs for the neural network. An example of the resulting τ_w distribution over a full pitching cycle is shown in Figure 2.

Neural Network Architecture and Training Details

Flow matching has emerged as a promising alternative to diffusion models due to its more direct and efficient approach to generative modeling. Flow matching holds the potential to produce high-quality samples with fewer function evaluations (Ref. 16). In this work, flow matching with conditional optimal transport is adopted to predict the flow field distributions, conditioning the process with the physical operating state.

Transformer architectures have demonstrated broad success across domains such as natural language processing, computer vision, reinforcement learning, and meta-learning (Ref. 17). Their scalability with respect to model size, data, and compute has led to state-of-the-art performance in both autoregressive modeling and vision tasks, including Vision Transformers (ViTs) (Ref. 18). In this work, we adopt the Diffusion Transformer (DiT) architecture proposed by (Ref. 19) as the backbone of our flow-matching model, with minor modifications. The model conditions its predictions via adaptive layer normalization, adaLN-Zero, where the scale and shift parameters are derived from a joint embedding of the time step t and operating conditions c. To improve computational efficiency, we implement several architectural changes. Quadratic attention is replaced with linear attention, and the original label embedder is substituted with a two-layer linear network. The first layer employs a SiLU activation function, and the second maps the result to the desired embedding of the airfoil's operating conditions. The model is also downsized to include 4 DiT blocks, a patch size of 4, a hidden size of 128, and 4 attention heads. This results in a compact model with 1.295 million parameters and a computational cost of 1.682 GFLOPs.

The implementation is based on PyTorch v2.5.1 and leverages the Distributed Data Parallel (DDP) module to scale training across devices. Neural network compilation is enabled to reduce overhead. The AdamW optimizer is used with a fixed learning rate of 10^{-4}, $\beta_1 = 0.5$, $\beta_2 = 0.999$, and zero weight decay. A global batch size of 64 is adopted to avoid the need for learning rate scaling. Training is conducted for 32 000 epochs (equivalent to 2 048 000 training iterations), and an Exponential Moving Average (EMA) with a decay factor of 0.999 is adopted.

To assess the physical fidelity of the predictions, we compute residuals associated with key physical laws. The ideal gas law residual, $\mathscr{R}_{\text{ig}}$, is calculated by comparing the predicted pressure P with the pressure inferred from the predicted density ρ and temperature T using the equation of state:

$$\begin{aligned} \mathscr{R}_{\text{ig}} &= P - P_{\text{ig}} \\ &= P - \rho R T \end{aligned}$$

Similarly, the skin friction residual, $\mathscr{R}_{\tau}$, quantifies the difference between the predicted skin friction and the value com-

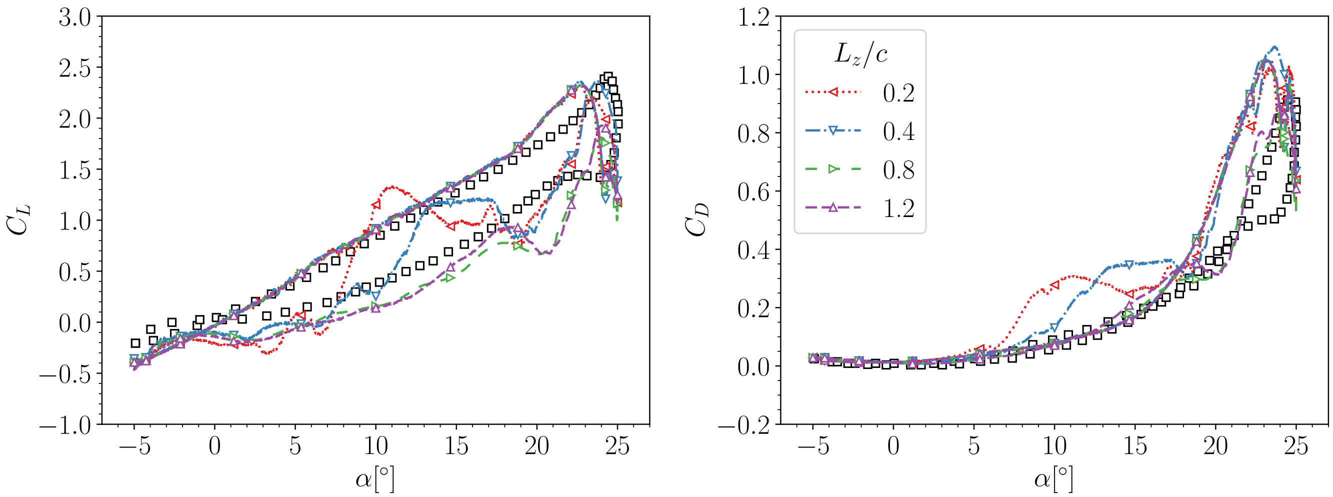

Figure 3. Spanwise size comparison of lift and drag coefficients for the simulated cycle. Experimental data from (Ref. 9).

puted from the tangential velocity gradient scaled by the local viscosity:

$$\begin{aligned}\mathscr{R}_\tau &= \tau_w - \tau_w^* \\ &= \tau_w - \mu \left(\frac{\partial u_t}{\partial n}\right)_{n=0} \\ &= \tau_w - \mu_0 \frac{T_0 + S}{T + S} \left(\frac{T}{T_0}\right)^{\frac{3}{2}} \sqrt{\left(\frac{\partial u_t}{\partial x}\right)_{n=0}^2 + \left(\frac{\partial u_t}{\partial y}\right)_{n=0}^2}\end{aligned}$$

where μ is the dynamic viscosity determined via Sutherland's law, and $\partial u_t/\partial n$ denotes the wall-normal gradient of the tangential velocity component.

RESULTS

WRLES analysis

This study investigates the influence of limited spanwise wing dimensions on dynamic stall behavior, employing wing geometries with four different span-to-chord ratios: $L_z/c = 0.2, 0.4, 0.8, 1.2$.

Figure 3 illustrates the instantaneous lift and drag coefficients as functions of the angle of attack for different span sizes. If the flow remains predominantly attached, all configurations exhibit similar aerodynamic behavior. During the early upstroke phase, from $\alpha = -5°$ to approximately $\alpha = 22°$, the variation in lift coefficient C_L across span-to-chord ratios remains within 0.05. However, significant divergence emerges as the LSB detaches and a DSV forms around $\alpha = 22°$, particularly for the two smallest span-to-chord ratios. Notably, these configurations exhibit a more pronounced first lift peak during the pitch-down phase, surpassing the corresponding peaks observed for larger spans. This difference is accentuated at lower angles of attack, where interaction between the DSV and trailing-edge vortex (TEV) yields a lift enhancement of up to 0.5 in C_L, and a drag increase of 0.3 in C_D during the climb phase. Conversely, larger-span configurations follow trends more consistent with experimental data, although with an approximate offset of 0.25 in C_L during pitch-down.

The temporal evolution of the DSV and TEV is depicted in Figure 4, showing spanwise-averaged distributions of pressure coefficient C_p and skin friction coefficient C_f on the suction side. Transition from a laminar to turbulent boundary layer begins at the trailing edge near $\alpha = 2°$, progressing toward the LSB by $\alpha = 10°$. By $\alpha = 15°$, the DSV becomes discernible, fully developing and reaching the trailing edge at $\alpha = 22°$, where it interacts with the TEV. Post-interaction flow behavior varies with spanwise extent. For $L_z/c = 0.2$ and 0.4, two dominant pressure peaks emerge at the trailing edge, followed by a weaker third, along with vortex shedding from the mid-chord. In contrast, larger-span geometries exhibit a distinct pattern: while the DSV–TEV interaction still induces a pressure peak, it is of lower magnitude and is followed by full flow separation without clear recurring structures.

Figure 5 shows contours of instantaneous velocity magnitude at selected angles of attack, providing further insight into span-to-chord effects on flow dynamics. During the upstroke, spanwise dimension has minimal influence; similar flow structures are observed across all cases. However, at the maximum angle of incidence, the impact of spanwise extent becomes evident. Smaller span-to-chord ratios (0.2 and 0.4) exhibit more coherent, two-dimensional DSV cores with elevated velocity peaks, a result of limited spanwise interaction due to boundary constraints. Larger-span configurations develop finer-scale turbulent structures, indicative of enhanced three-dimensional interactions. This distinction intensifies during the descent phase. Configurations with smaller spans display persistent, pronounced vortical structures, while larger spans show a more diffuse, fully stalled wake. A particularly notable contrast arises at $\alpha = 10°$, where the largest-span geometry shows complete stall, while the $0.8c$ case retains oscillatory wake patterns.

Further evidence of three-dimensionality is provided by the spanwise distributions of C_f, shown in Figure 6 at $\alpha = 10°$

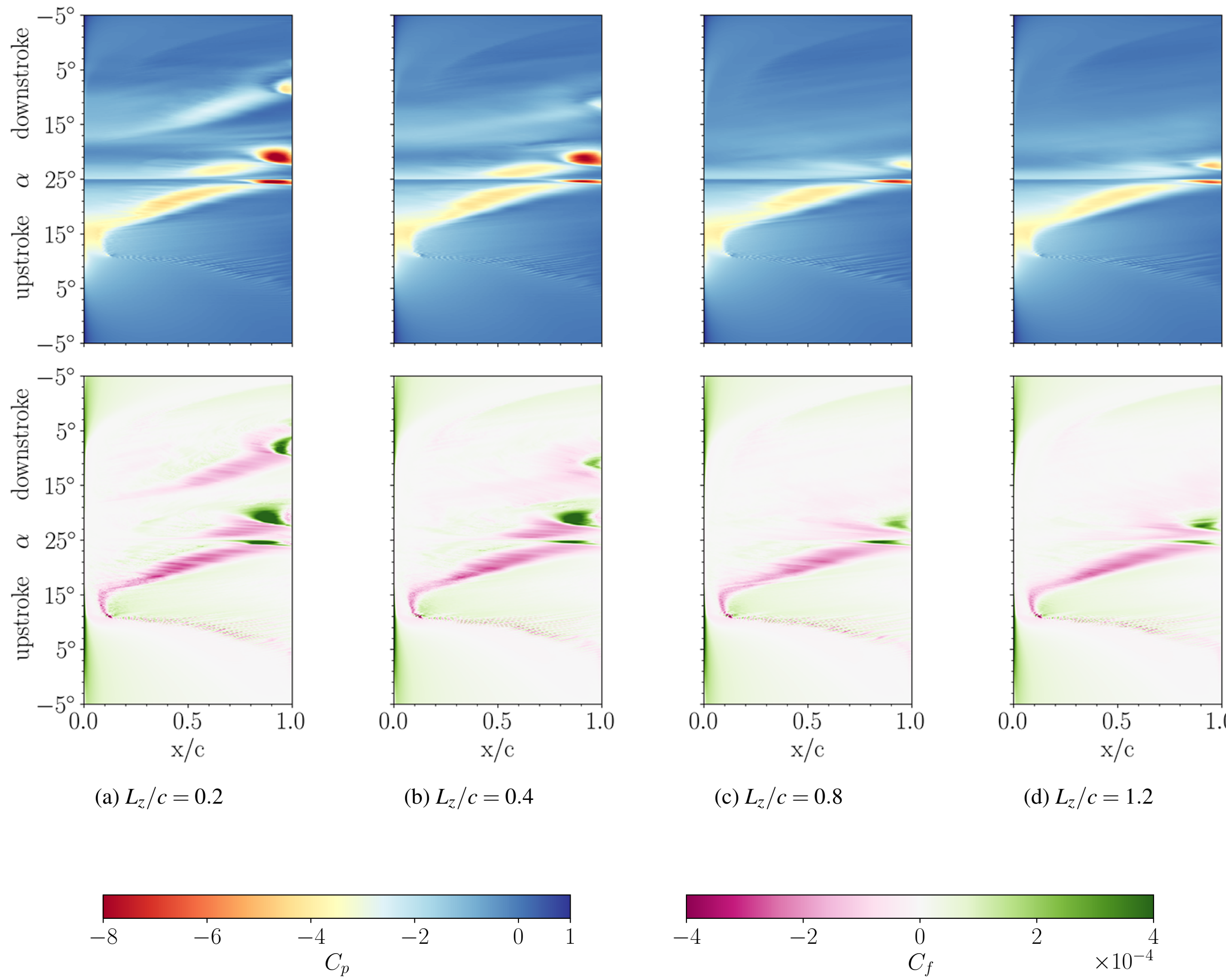

Figure 4. Spanwise averaged pressure and friction coefficient over the suction side of the wing section.

and $\alpha = 22°$, covering both upstroke and downstroke phases. These contours reinforce prior observations: during upstroke, recirculation zones, identified by regions of negative C_f, remain similarly aligned across all cases. The most notable variations arise at the maximum angle of attack and during descent, where three-dimensional structures emerge prominently for larger-span geometries. Interestingly, at $\alpha = 22°$, the C_f distributions converge across all configurations, suggesting that at peak incidence, the spanwise effects become less pronounced. Additionally, at $\alpha = 10°$, the two largest-span cases exhibit comparable qualitative behavior, further emphasizing the role of geometric extent in modulating transitional flow regimes.

Comparison with RANS and hybrid RANS/LES approaches

We now present a comparative analysis between the LES results discussed in the previous section and those reported by (Ref. 3). In their study, two different numerical approaches were employed: a two-dimensional unsteady RANS simulation and a three-dimensional hybrid RANS/LES method. For the RANS simulation, the k–ω SST turbulence model was used in combination with the γ-intermittency transition model, referred to as iSST, to more accurately capture the laminar-to-turbulent transition within the boundary layer. The hybrid approach utilized a Stress-Blended Eddy Simulation (SBES) framework, also coupled with the same RANS turbulence model. The LES corresponds to the geometry with a span-to-chord ratio of 1.2.

Figure 7 compares the profile loads obtained from these numerical simulations with experimental measurements reported by (Ref. 9). For attached flow conditions, the predicted C_L curves from all numerical models are closely aligned, with a maximum discrepancy of 0.05. The most significant deviation between numerical and experimental results occurs in the angle of attack corresponding to peak lift: 22° for the simulations versus 24° in the experimental data. Interestingly, all numerical models, including the WRLES approach, con-

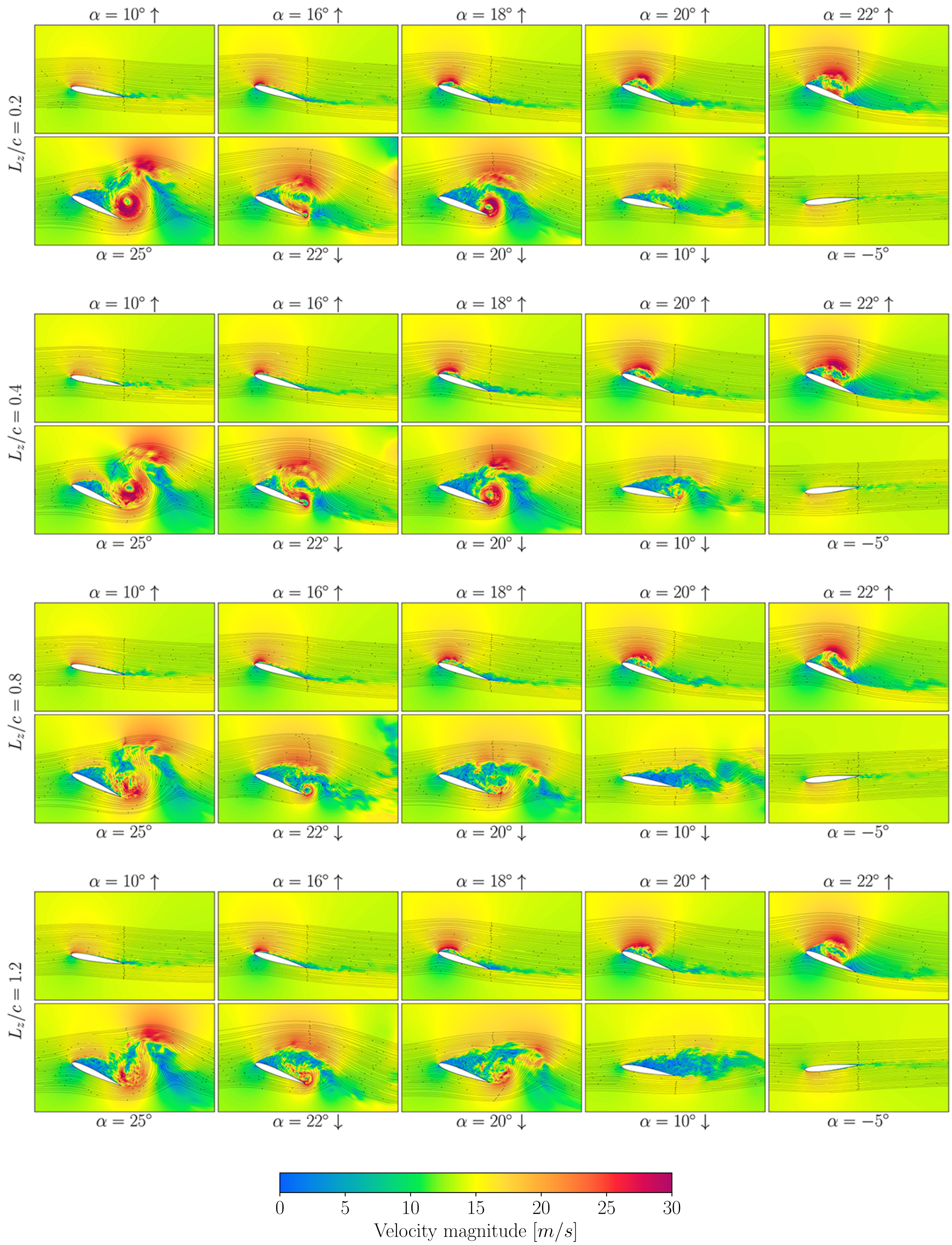

Figure 5. Instantaneous velocity magnitude contour at mid-section as a function of span-to-chord ratio. ↑ indicates the upstroke phase while ↓ the downstroke one.

sistently predict an earlier downward convection of the DSV compared to the experiments. In detail, the SBES model underpredicts the peak C_L by approximately 0.2 compared to the LES solution. In contrast, the iSST model provides a comparable prediction, exhibiting a phase lag of approximately 0.5 degrees. A similar pattern is observed in the C_D, with SBES

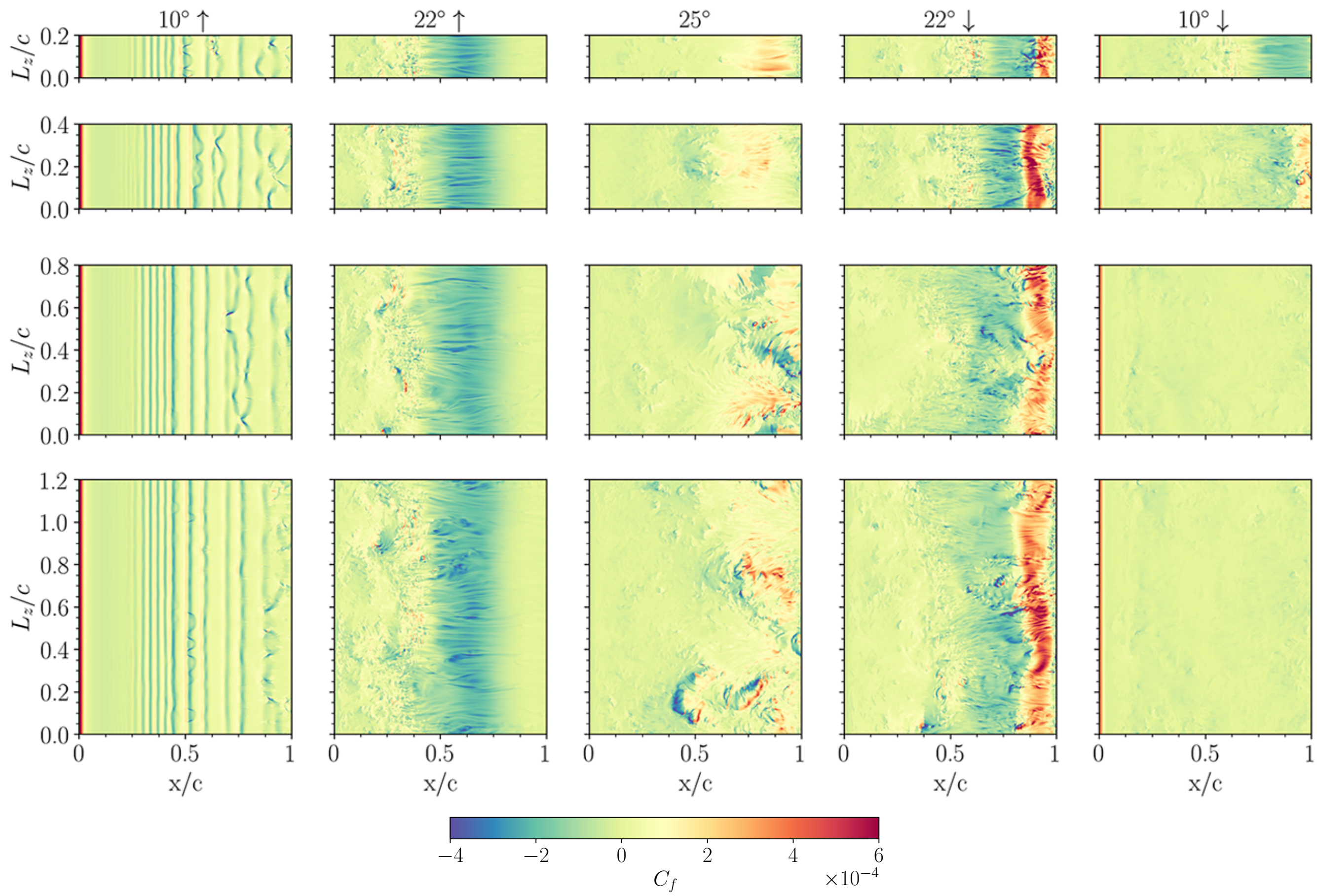

Figure 6. Instantaneous friction coefficient contour over the suction side of the airfoil as a function of span-to-chord ratio. ↑ indicates the upstroke phase while ↓ the downstroke one.

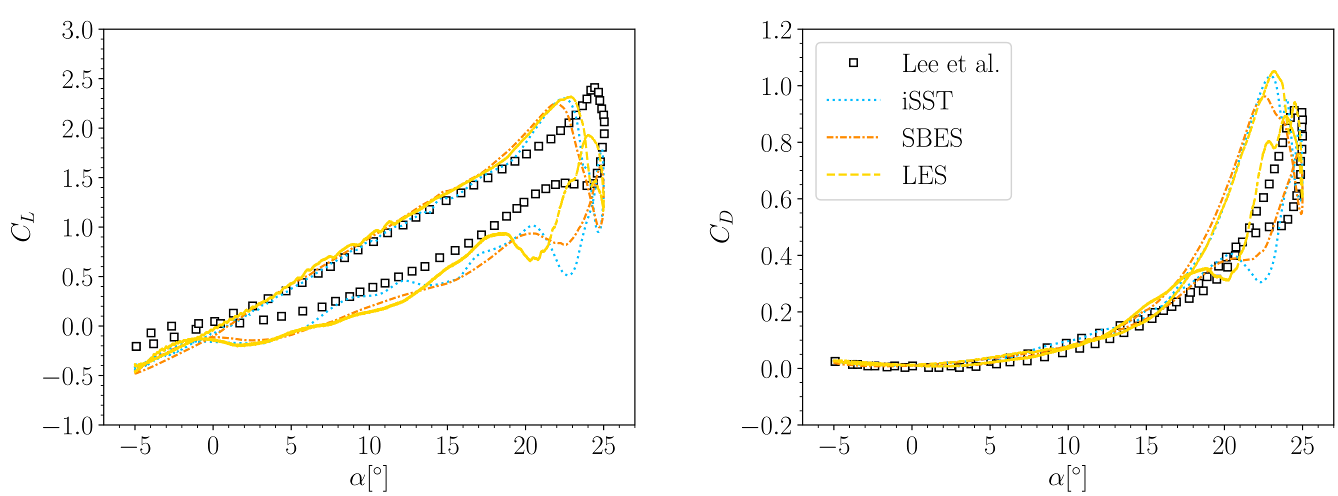

Figure 7. Numerical simulation comparison of lift and drag coefficients. LES refers to the 1.2*c* spanwise case. SBES and iSST data from (Ref. 3). Experimental data from (Ref. 9).

underestimating the peak value by around 0.1 compared to the LES. During the initial descent phase of the motion, $\alpha = 24°$ to 18°, which is characterized by fully stalled flow, the models begin to diverge. However, as the angle of incidence continues to decrease, the results are closer. It is important to note that cycle-to-cycle variations in the stalled regime are not considered due to the high computational cost associated with LES. Despite a persistent offset in C_L of approximately 0.2 compared to the experimental data, the overall simulation trends show good agreement with experimental observations.

Deep learning surrogate model

The predictive capability of the proposed surrogate model is evaluated for unsteady flow fields over an airfoil undergoing pitching motion. Model performance is assessed through direct comparison with iSST RANS, with particular emphasis on the model's ability to generalize to previously unseen flow conditions and to maintain accuracy across a range of operating regimes. The evaluation begins with a qualitative analysis, wherein the spatial statistics, namely the mean and standard deviation, of the predicted flow fields are compared against those obtained from the reference simulations. This is followed by a quantitative assessment in which these statistical measures are examined as functions of the FM time steps. To further characterize the model's performance, the distributions of the predicted flow fields are analyzed. Specifically, we visualize the temporal evolution of these distributions and compute the Wasserstein distance between predicted and reference fields for six representative flow variables over the FM time horizon. Moreover, the physical consistency of the model is investigated by quantitatively comparing two physical residuals as functions of FM time steps. Lastly, the computational efficiency of the surrogate model is briefly discussed by reporting the inference time required to generate flow field predictions.

Figure 8 provides a comparative analysis of the predicted fluid fields against the reference data from high-fidelity CFD simulations. The mean fields show strong agreement with the reference data, particularly at the extrema, where the predicted values closely match the ground truth. For the standard deviation, the predictions effectively capture the overall distribution of numerical fluctuations. However, minor discrepancies appear in peak values, notably in the velocity gradient along the x-direction and temperature. Despite these deviations, the predictive model demonstrates strong reliability in representing the underlying flow characteristics.

A more detailed quantitative analysis is presented in Figure 9, which shows the mean squared error (MSE) for both the mean and standard deviation as a function of the number of integration steps used in flow matching. The results indicate a rapid decrease in error from 1 to 10 time steps, after which it stabilizes. The optimal number of neural network evaluations for peak performance falls within the range of 10 to 20.

Further evaluation is presented in Figure 10, which depicts the results using 20 FM time steps, and Figure 11, which illustrates the Wasserstein distance as a function of the number of neural network evaluations. The histograms show that the reference and predicted data are nearly identical. However, slight discrepancies are present, particularly in the temperature field, where minimum values slightly deviate from 230K. The second figure quantitatively assesses model differences based on optimal transport theory. With a single time step, the error is approximately $1.5\times$ higher, but from 2 time steps onward, no significant variations are observed, demonstrating the robustness of the proposed method.

Finally, we assess the physical residuals derived from the predicted data. Figure 12 presents residuals associated with the ideal gas law and skin friction computation. A key observation is that the mean error is in the order of $\sim 10^{-5}$ confirming that the reduced order model is able to reproduce the underlying physics.

Inference performance

Although direct one-to-one comparisons between CFD simulations and neural network predictions are inherently challenging, we report inference times for our proposed framework on both CPU and GPU platforms to provide context for performance. Specifically, the hardware used includes a 32-core CPU and an NVIDIA A100 GPU. For CPU-based inference, we limited usage to 8 cores, matching the resources allocated per CFD simulation run. Under this configuration, a full CFD simulation requires, on average, 76 minutes to complete. Table 4 summarizes the complete runtime comparison. In contrast, the surrogate model running on the GPU can generate predictions for 128 samples in approximately 0.1 seconds, offering a significantly faster alternative for real-time load estimation in aeroelastic analyses and digital twin applications.

CONCLUSIONS

- The study confirms that the span-to-chord ratio plays a pivotal role in accurately describing dynamic stall. Simulations with limited spanwise extents display exaggerated two-dimensional flow features, leading to overestimated lift peaks and coherent vortical structures. In contrast, larger spans introduce the necessary three-dimensional effects that align more closely with experimental observations, especially in the development and breakdown of the DSV.

- WRLES provided deep insights into the dynamic stall process, resolving complex transitional and turbulent flow features across the airfoil surface. The simulation data served as a high-fidelity reference for assessing the performance of traditional turbulence models used in RANS and hybrid RANS/LES approaches.

- A diffusion-based deep learning surrogate, built on a modified DiT architecture, was trained using flow-matching techniques. The model demonstrated excellent capability to predict both statistical and instantaneous fields across a wide range of operating conditions. It generalized well beyond the training data and maintained physical accuracy, as confirmed by residual analysis on key conservation laws.

- The reduced-order model drastically reduced computational cost compared to RANS simulations. With inference times under a second for batch predictions, the model is well-suited for integration into real-time control systems, aeroelastic simulations, and digital twin frameworks for rotorcraft.

Author contact: Giacomo Baldan giacomo.baldan@polimi.it

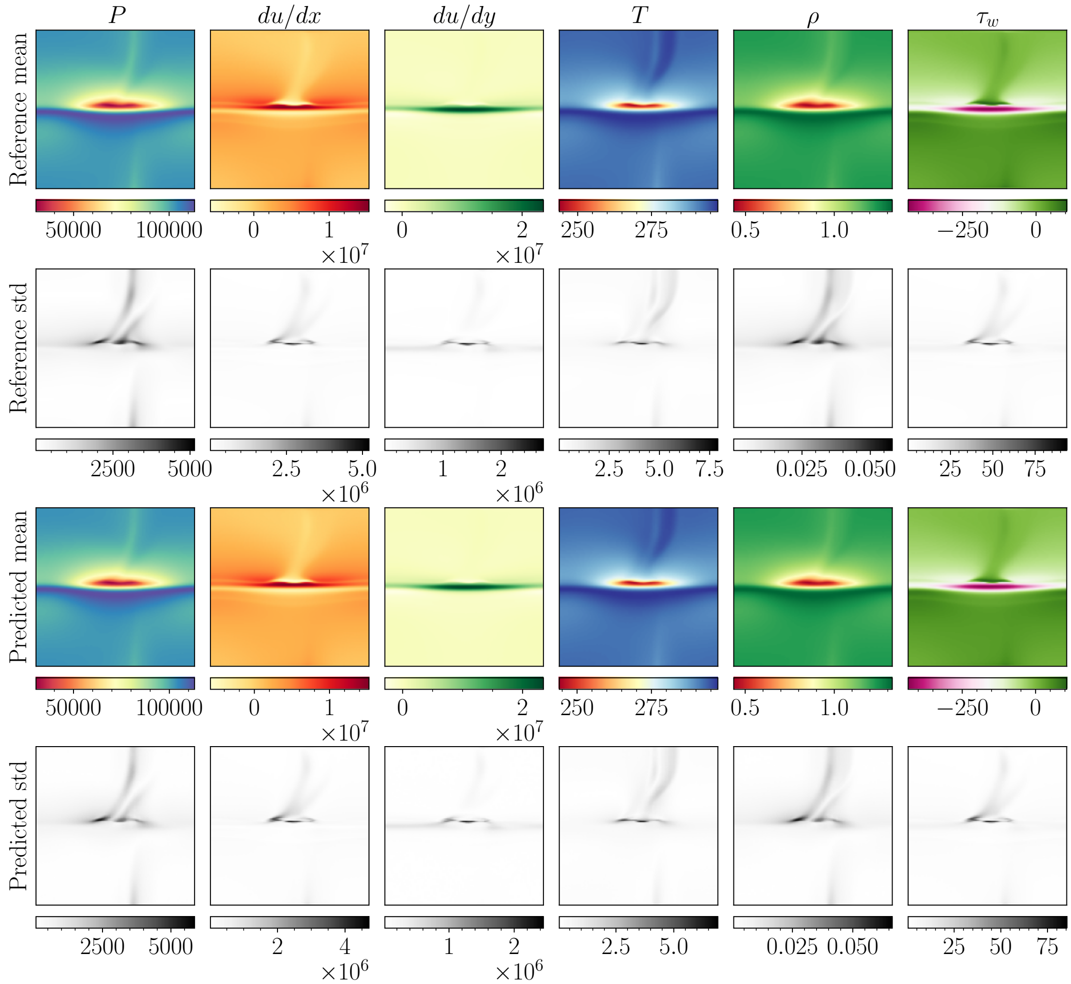

Figure 8. Example of mean and standard deviation for reference and predicted spatio-temporal fields. The number of flow matching steps is 20.

Table 4. Wall-clock time in seconds to generate n samples using 8 CPU cores and an NVIDIA A100. Batch size is set equal to the number of samples. One simulation with the same cores takes on average 76 minutes ($4.56 \cdot 10^3$ s).

FM	Number of samples - GPU			Number of samples - CPU		
steps	**32**	**64**	**128**	**32**	**64**	**128**
1	$3.23 \cdot 10^{-3}$	$3.38 \cdot 10^{-3}$	$3.80 \cdot 10^{-3}$	$2.27 \cdot 10^{0}$	$6.16 \cdot 10^{0}$	$1.20 \cdot 10^{1}$
2	$9.76 \cdot 10^{-3}$	$1.50 \cdot 10^{-2}$	$2.53 \cdot 10^{-2}$	$4.48 \cdot 10^{0}$	$1.11 \cdot 10^{1}$	$2.40 \cdot 10^{1}$
5	$2.90 \cdot 10^{-2}$	$4.95 \cdot 10^{-2}$	$8.99 \cdot 10^{-2}$	$1.17 \cdot 10^{1}$	$2.96 \cdot 10^{1}$	$5.99 \cdot 10^{1}$
10	$6.15 \cdot 10^{-2}$	$1.07 \cdot 10^{-1}$	$1.98 \cdot 10^{-1}$	$2.25 \cdot 10^{1}$	$5.09 \cdot 10^{1}$	$1.19 \cdot 10^{2}$
20	$1.26 \cdot 10^{-1}$	$2.23 \cdot 10^{-1}$	$4.13 \cdot 10^{-1}$	$4.68 \cdot 10^{1}$	$1.15 \cdot 10^{2}$	$2.40 \cdot 10^{2}$
50	$3.21 \cdot 10^{-1}$	$5.70 \cdot 10^{-1}$	$1.06 \cdot 10^{0}$	$1.12 \cdot 10^{2}$	$3.08 \cdot 10^{2}$	$5.99 \cdot 10^{2}$

ACKNOWLEDGMENTS

The authors gratefully acknowledge Leonardo SpA–Helicopter Division for providing access to the `davinci-1` supercomputer, on which the computations were performed.

REFERENCES

1. Baldan, G., Heuschneider, V., Mitropoulos, M., Yavrucuk, I., and Guardone, A., "Numerical and Experimental Investigation of Stall on the MERIT Rotor in Hover," *The Vertical Flight Society's 80th Annual Forum & Technology Display, May 7-9*, 2024. DOI: 10.4050/F-0080-2024-1118.

2. Gardner, A. D., Jones, A. R., Mulleners, K., Naughton, J. W., and Smith, M. J., "Review of rotating wing dynamic stall: Experiments and flow control," *Progress in Aerospace Sciences*, Vol. 137, 2023, pp. 100887.

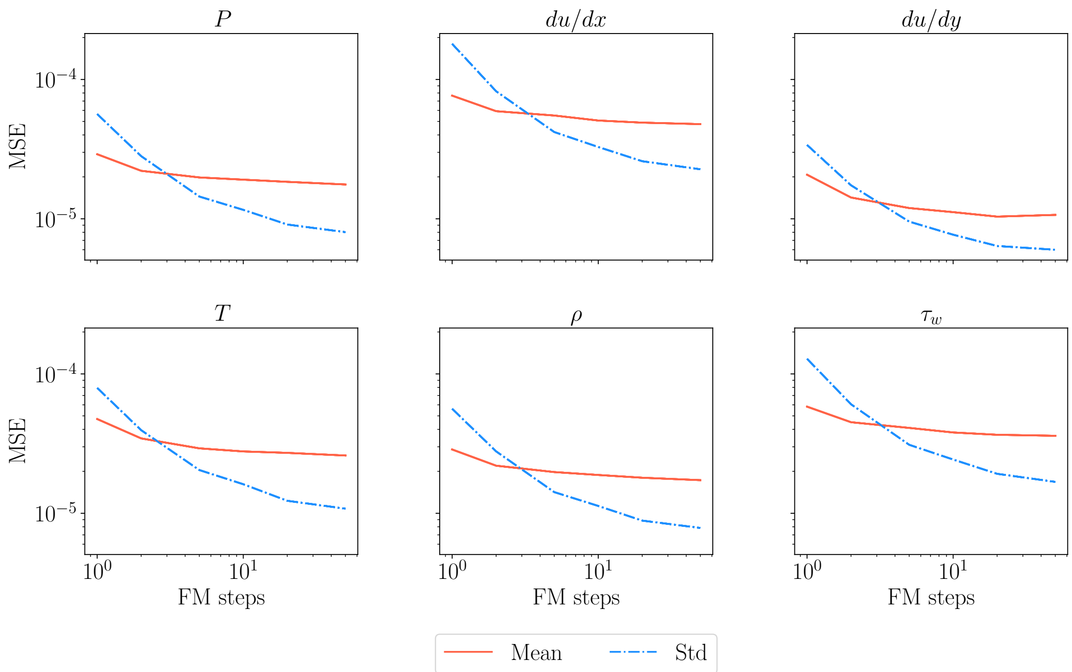

Figure 9. Mean squared error for predicted field mean and standard deviation.

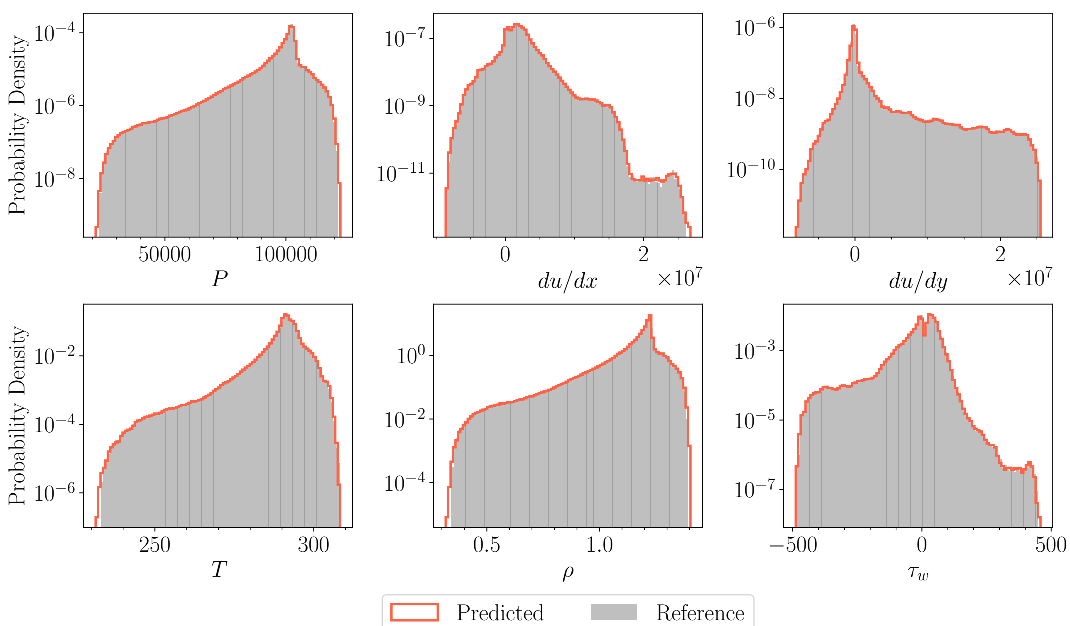

Figure 10. Comparison of predicted field distributions. The number of flow matching steps is 20.

DOI: 10.1016/j.paerosci.2023.100887.

3. Baldan, G., Manara, F., Frassoldati, G., and Guardone, A., "The effects of turbulence modeling on dynamic stall," *Acta Mechanica*, Vol. 236, (2), Feb 2025, pp. 1411–1427. DOI: 10.1007/s00707-025-04232-w.

4. Shur, M., Spalart, P., Strelets, M., and Travin, A., "Detached-eddy simulation of an airfoil at high angle of attack," *Engineering Turbulence Modelling and Experiments 4*, Elsevier Science Ltd, Oxford, 1999, pp. 669–678. DOI: 10.1016/B978-008043328-8/50064-3.

5. McCroskey, W., *The Phenomenon of Dynamic Stall*, Defense Technical Information Center, 1981.

6. Baldan, G., and Guardone, A., "Wall-resolved large eddy simulations of a pitching airfoil in deep dynamic stall," *Physics of Fluids*, Vol. 37, (2), 02 2025, pp. 024112. DOI: 10.1063/5.0252828.

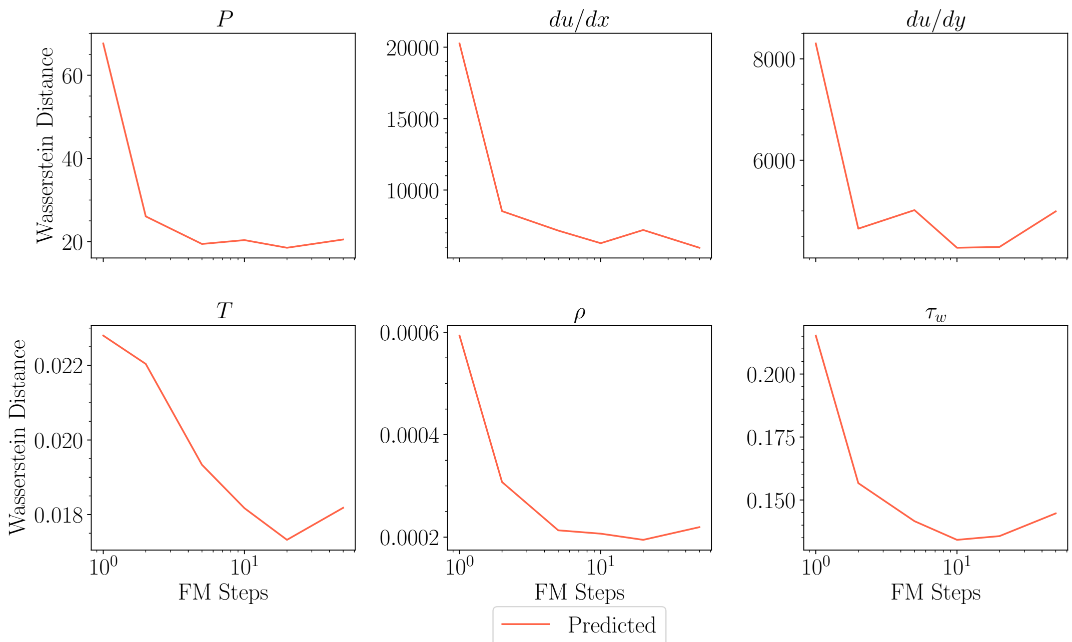

Figure 11. Wasserstein distance of predicted field distributions.

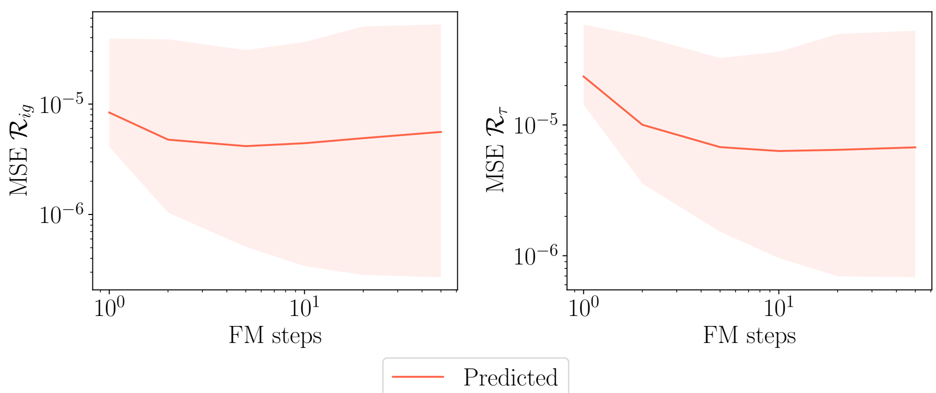

Figure 12. Mean squared error of the predicted field residuals. Band refers to the maximum and minimum error.

7. Beddoes, T. S., "A third generation model for unsteady aerodynamics and dynamic stall," Technical Report RP-908, 1993.

8. Truong, V. K., "Prediction of helicopter rotor airloads based on physical modeling of 3-D unsteady aerodynamics," 22nd European Rotorcraft Forum, 1996.

9. Lee, T., and Gerontakos, P., "Investigation of flow over an oscillating airfoil," *Journal of Fluid Mechanics*, Vol. 512, 2004, pp. 313–341. DOI: 10.1017/S0022112004009851.

10. Piomelli, U., and Balaras, E., "Wall-layer models for large-eddy simulations," *Annual Review of Fluid Mechanics*, Vol. 34, (1), 2002, pp. 349–374. DOI: 10.1146/annurev.fluid.34.082901.144919.

11. Georgiadis, N. J., Rizzetta, D. P., and Fureby, C., "Large-eddy simulation: current capabilities, recommended practices, and future research," *AIAA Journal*, Vol. 48, (8), 2010, pp. 1772–1784. DOI: 10.2514/1.J050232.

12. Choi, H., and Moin, P., "Grid-point requirements for large eddy simulation: Chapman's estimates revisited," *Physics of Fluids*, Vol. 24, (1), 01 2012, pp. 011702. DOI: 10.1063/1.3676783.

13. Baldan, G., and Guardone, A., "Pattern recognition of the flow around a pitching NACA 0012 airfoil in dynamic stall conditions," Aeronautics and Astronautics, 2023. DOI: 10.21741/9781644902813-35.

14. Baldan, G., Guardone, A., and Thuerey, N., "Flow Matching for Stochastic Reduced Order Modeling of

Unsteady Flows over Pitching Airfoils," *Computers & Fluids*, 2025.

15. Kocis, L., and Whiten, W. J., "Computational Investigations of Low-Discrepancy Sequences," *ACM Trans. Math. Softw.*, Vol. 23, (2), jun 1997, pp. 266–294. DOI: 10.1145/264029.264064.

16. Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nickel, M., and Le, M., "Flow Matching for Generative Modeling," The Eleventh International Conference on Learning Representations, 2023.

17. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I., "Attention Is All You Need," , 2023.

18. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," , 2021.

19. Peebles, W., and Xie, S., "Scalable Diffusion Models with Transformers," , 2023.