SURROGATE MODELING OF HIGHLY FLEXIBLE STRUCTURES AND AERODYNAMICS USING NEURAL NETWORKS

Vitor B. Santos[1], Breno S. C. Vieira[1], Flávio L. Cardoso-Ribeiro[1], Antônio B. Guimarães Neto[1]

[1]Instituto Tecnológico de Aeronáutica
Pç. Mal. Eduardo Gomes 50, 12228-000 São José dos Campos, Brazil
vitor.santos@ga.ita.br
flaviocr@ita.br

Abstract: The renaissance of neural networks in the scientific community in recent years has brought new perspectives for improving the computational efficiency of traditional modeling techniques. Hamiltonian neural networks leverage the energy-preserving properties of the Hamiltonian formalism to provide surrogate models with increased interpretability compared to conventional feed-forward models. In this study, we employ a lumped-mass multibody method to derive the equations of motion of two highly flexible structures. We perform a model order reduction via modal decomposition while preserving the nonlinearities with the use of exact kinematic relations. After validating full- and reduced-order models, we use them to produce datasets and train the neural networks, which serve as ready-to-use surrogate models. Preliminary findings show that the surrogate models based on neural networks can significantly reduce the time necessary to simulate the free response of the structures. Furthermore, we demonstrate that surrogate models based on Hamiltonian neural networks have energy-preserving capabilities, maintaining accuracy levels even for long simulations. Due to their architecture, when external loads are considered, the surrogate models require the analytical calculation of the generalized forces, jeopardizing the efficiency gains obtained by our approach. We also present initial findings on the use of neural networks for faster aerodynamic models for flexible aircraft, particularly as surrogate models for the vortex-lattice method. By using a neural network as the aerodynamic surrogate model in a specific flexible aircraft simulation framework, the computational costs were reduced by a factor of 100 on average. The outcomes of this study demonstrate that surrogate models based on neural networks can soon become an efficient and reliable alternative for modeling arbitrarily flexible aircraft, provided the current limitations are addressed.

1 INTRODUCTION

As the world faces climate change, the aeronautical industry is developing and improving solutions aimed at reducing the environmental impact of aviation. Since the late 1960s, the aspect ratio of lifting surfaces has increased and become an important asset for enhancing aerodynamic performance, as illustrated in Figure 1. New aircraft designs combine the use of modern materials and structural optimization to improve resistance while reducing structural weight. These advancements allow for higher aspect ratio lifting surfaces, significantly reducing drag, which leads to reduced fuel consumption and increased flight range.
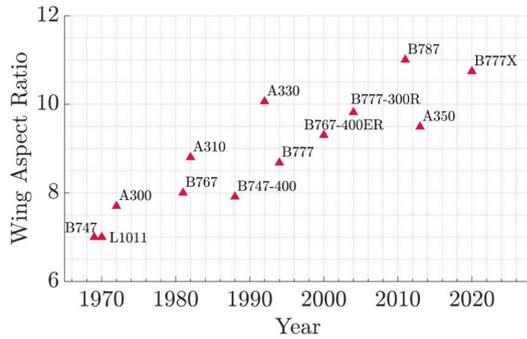
Figure 1: Aspect ratio trends for commercial aircraft wings since the end of 1960s [1].
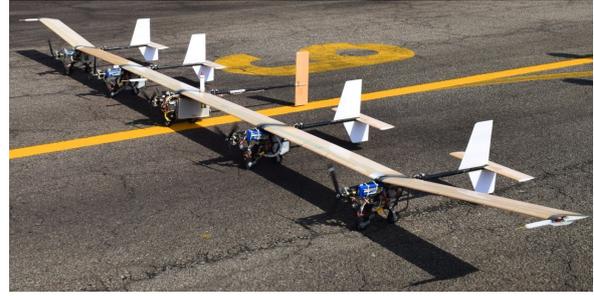

Figure 2: X-HALE-BR prototype, a highly flexible aircraft [2].

Despite their advantages, high aspect ratio lifting surfaces suffer from a serious drawback: high structural flexibility. In addition to material and geometrical nonlinearities, highly flexible structures aggravate the coupling between structure and airflow dynamics. Moreover, it is often necessary to consider unsteady aerodynamic models and the effects of high angles of attack [3]. Some mathematical models for highly flexible aircraft (such as the one depicted in Figure 2) have been derived so far (see [4, 5]), however, these models are computationally demanding and impractical for real-time simulations, posing challenges in design optimization and control projects.

The recent resurgence of neural networks, enabled by improvements in computational power, has recently provided outstanding solutions for benchmark problems in various fields. In exact sciences, this trend has been fueled by the emergence of scientific machine learning [6], where traditional machine learning is combined with physically-based modeling techniques to obtain efficient and reliable surrogate models. Particularly, some interesting scientific machine learning models leverage energy-preservation principles, such as Lagrangian neural networks [7], Hamiltonian neural networks [8–10], and port-Hamiltonian neural networks [11].

This study investigates how Hamiltonian neural networks (HNNs) can be used as surrogate models for two highly flexible structures and how they perform compared to a reduced-order reference model derived from the modal decomposition of a full-order model based on a lumped-mass multibody approach. We examine whether incorporating the underlying physics provided by HNNs offers improvements over a baseline surrogate model obtained from a conventional feed-forward neural network, also known as a multi-layer perceptron (MLP). As our future goal is to model flexible and highly flexible aircraft, we also offer preliminary insights into employing neural networks to accelerate low-fidelity aerodynamic models, particularly those based on the vortex-lattice method.

The text is organized such that in Section 2 we derive the full- and reduced-order reference models for highly flexible slender beams. In Section 3 we describe the architecture of the Hamiltonian neural networks and how they differ from conventional feed-forward neural networks. We validate our reference models in Section 4, where we also compare the efficiency and accuracy of the surrogate models by simulating a highly flexible beam and a highly flexible beam-like wing under various free and forced conditions. Section 5 presents perspectives and preliminary outcomes of using neural networks as replacements for the vortex-lattice method. Finally, in Section 6, we conclude our work, summarizing the main outcomes and limitations found, and presenting interesting perspectives for future research.

The following conventions are assumed throughout this text: matrices are written in uppercase bold letters (e.g. $\mathbf{A}$), vectors in lowercase bold letters (e.g. $\mathbf{a}$), scalars in italic uppercase or lowercase letters (e.g. A or a), time derivatives are represented using Newton's notation (e.g. $\dot{a}$), and functions are denoted by the symbolic letter f.

2 THE REFERENCE HIGHLY FLEXIBLE BEAM MODEL

We derive the reference structural model used in this study from a straightforward methodology presented in [12] and [13]. Using the principles of rigid multibody dynamics combined with a lumped mass distribution and Hamiltonian mechanics, a highly flexible cantilever beam can be discretized into a number N of rigid elements. These elements are connected to the clamping point through ideal joints: rotary joints are used for out-of-plane bending, in-plane bending, and torsion deformations, while linear joints are used for axial deformation.

Figure 3 illustrates how the discretization scheme works for out-of-plane bending deformation. The discretization of the other deformation modes follows analogously. The discretization mentioned implies that any joint i can be associated with the generalized coordinate q_i.

The number of degrees-of-freedom n of the system is given by the number of elements of the discretized beam multiplied by the number of deformation modes $n_d = \{n_d \in \mathbb{N} : n_d \leq 4\}$ considered. In this study, we assume that only out-of-plane bending and torsion deformation modes are relevant, therefore $n = 2N$.

The kinematic relations of the structure require 2N+1 orthonormal reference frames defined in Cartesian coordinate systems: (i) one global reference frame G, inertial and with origin at the clamping point; (ii) N local undeformed frames U_i, inertial and defining at the origin of each i-th undeformed element; and (iii) N local deformed frames D_i, defined at the origin of each i-th deformed element. Figure 4 illustrates the different reference frames used.
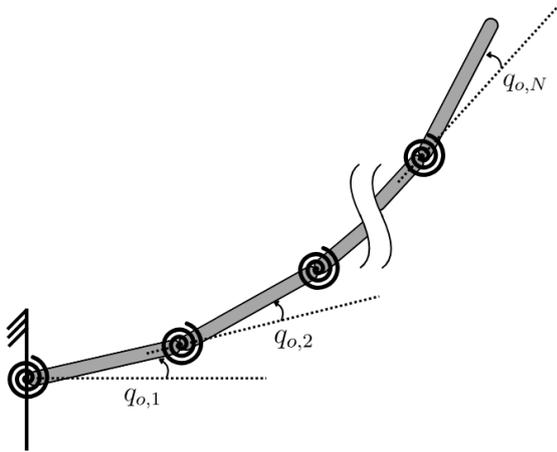

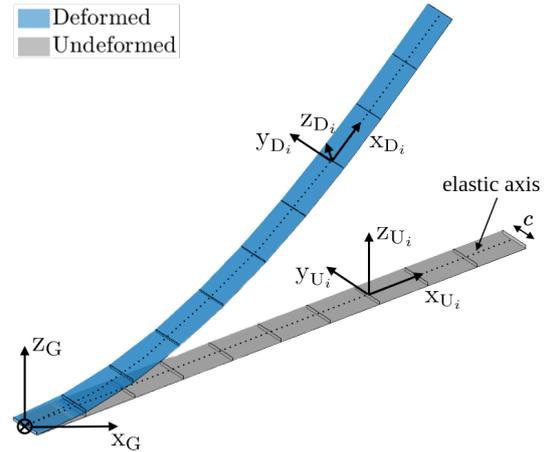Figure 3: Beam discretization for out-of-plane bending deformation.

Figure 4: The reference frames used to describe the position of the elements in the structure.

The position vector $\mathbf{r}$ of an arbitrary point P at element i in the global frame is given by:

$$\mathbf{r}_{P,i} = \mathbf{C}_{G/U_i} \mathbf{C}_{U_i/D_i} \mathbf{r}_{P,i,D_i}, \tag{1}$$

where $\mathbf{C}_{G/U_i} := f(q_{1,1}, q_{1,2}, ..., q_{k,i})$ is a function of all generalized coordinates associated with all k degrees-of-freedom from the first until the i-th element. In other words, it is the rotation matrix from the local undeformed frame to the global inertial frame. $\mathbf{C}_{U_i/D_i} := f(\psi, \theta, \phi)$ is the

rotation matrix from the local deformed frame to the local undeformed frame, and it is defined by the angles of rotation ψ, θ and ϕ of the undeformed structure relative to the global reference frame, assuming the 3-2-1 Euler angles convention.

In the Hamiltonian formalism, the equations of motion for the k-th degree-of-freedom of a holonomic system subjected to conservative, nonconservative, and dissipative forces, assume the following form:

$$
\begin{aligned}
\dot{q}_k &= \frac{\partial \mathcal{H}}{\partial p_k}, \\
\dot{p}_k &= -\frac{\partial \mathcal{H}}{\partial q_k} - \frac{\partial \mathcal{R}}{\partial \dot{q}_k} + Q_k,
\end{aligned}
\tag{2}
$$

where q_k and p_k are the generalized coordinate and momentum of the k-th degree-of-freedom, respectively. $\mathcal{H} := f(\mathbf{q}, \mathbf{p})$ is the Hamiltonian function, $\mathcal{R} := f(\dot{\mathbf{q}})$ is the Rayleigh dissipation function, and $Q_k := f(t, \mathbf{q})$ corresponds to the generalized forces caused by external forces acting upon the k-th degree-of-freedom.

2.1 Full-Order Model

The Hamiltonian function can be obtained by the sum of kinetic and potential energies, as it is in our system. Luckily, in our system the kinetic energy T is not explicitly dependent on time, being given simply by a quadratic term on the generalized velocities, i.e.,

$$
T = \frac{1}{2} \dot{\mathbf{q}}^\top \left(\mathbf{J}^\top \mathbf{M} \, \mathbf{J} + \mathbf{I} \right) \dot{\mathbf{q}},
\tag{3}
$$

where $\mathbf{M}$ is the mass matrix and $\mathbf{J} := f(\mathbf{q})$ is the Jacobian matrix relative to the center of mass position of each element with respect to the generalized coordinates. The Jacobian matrix can be calculated using a numerical differentiation scheme, for instance, the complex step method [14]. The mass moment of inertia matrix $\mathbf{I}$ is non-zero only if we assume torsion deformations.

Even though the gravitational forces arise from a conservative potential, we decided to include its contribution in the generalized forces term. Therefore, the potential energy of the system is simply due to the elastic potential of the joints. Hence, the potential energy U of the system can be expressed by:

$$
U = \frac{1}{2} \mathbf{q}^\top \mathbf{K} \, \mathbf{q},
\tag{4}
$$

where $\mathbf{K}$ represents the stiffness matrix composed of the stiffness coefficients of the joints associated with each degree of freedom.

The Rayleigh dissipation function $\mathcal{R}$ represents velocity-dependent dissipative forces in a generalized form, convenient to the Hamiltonian formalism. It can be written as:

$$
\mathcal{R} = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{C} \, \dot{\mathbf{q}},
\tag{5}
$$

where $\mathbf{C}$ defines the damping matrix. In this study, we assume damping coefficients proportional to the structural stiffness by factors β_k, therefore $\mathbf{C} = \mathbf{K}\boldsymbol{\beta}$.

After some manipulation, Equation 2 can be arranged in the following matrix form:

$$
\begin{aligned}
\dot{\mathbf{q}} &= \left(\mathbf{J}^\top \mathbf{M} \, \mathbf{J} + \mathbf{I} \right)^{-1} \mathbf{p}, \\
\dot{\mathbf{p}} &= -\mathbf{K} \, \mathbf{q} - \mathbf{C} \, \dot{\mathbf{q}} - \frac{\partial}{\partial \mathbf{q}} \left[\mathbf{p}^\top \left(\mathbf{J}^\top \mathbf{M} \, \mathbf{J} + \mathbf{I} \right)^{-1} \mathbf{p} \right] + \mathbf{Q}.
\end{aligned}
\tag{6}
$$

A derivative term concerning $\mathbf{q}$ appears in Equation 6 and is related to gyroscopic effects. This term is computationally demanding and its effect on the solution is usually irrelevant, as demonstrated by [12] for a similar system. Henceforth, when mentioned, the term *reference full-order model* (or simply *full-order model* - FOM) refers to the following set of equations:

$$\dot{\mathbf{q}} = \left(\mathbf{J}^{\top}\mathbf{M}\,\mathbf{J} + \mathbf{I}\right)^{-1}\mathbf{p},$$
$$\dot{\mathbf{p}} = -\,\mathbf{K}\,\mathbf{q} - \mathbf{C}\,\dot{\mathbf{q}} + \mathbf{Q}. \tag{7}$$

It is important to mention that the nonlinearities in the equations of motion 7 are present since the Jacobian is a function of the generalized coordinates and computed using the exact kinematic relations demonstrated by Equation 1.

2.2 Reduced-Order Model

To derive a reduced-order model, we linearize the equations of motion 7 around an equilibrium position. The equilibrium solution can be calculated by setting all time-dependent terms to zero, as follows:

$$\mathbf{K}\mathbf{q} - \mathbf{Q} = \mathbf{0}. \tag{8}$$

The solution of the system of equations 8 can be found using an iterative scheme such as the Newton-Raphson method.

Once the equilibrium solution $\mathbf{q}_{\mathrm{eq}}$ is known, we can calculate the eigenvalues $\boldsymbol{\lambda}$ and modal matrix $\boldsymbol{\Phi}$ from the following equation:

$$\left[\mathbf{K} - \boldsymbol{\lambda}\left(\mathbf{J}_{\mathrm{eq}}^{\top}\,\mathbf{M}\,\mathbf{J}_{\mathrm{eq}} + \mathbf{I}\right)\right]\boldsymbol{\Phi} = \mathbf{0}, \tag{9}$$

where $\mathbf{J}_{\mathrm{eq}}$ is the Jacobian matrix at equilibrium condition, i.e. for $\mathbf{q} = \mathbf{q}_{\mathrm{eq}}$.

The eigenvectors (columns) of the modal matrix $\boldsymbol{\Phi} \in \mathbb{C}^{n \times n}$ calculated in Equation 9 can be reordered for a convenient arrangement of the eigenvalues, e.g. in descending order if we are interested in the low-frequency modes. We can truncate the modal matrix, such that the reduced modal matrix is $\boldsymbol{\Phi}_r \in \mathbb{C}^{n \times r}$, in which only the first r modes (lowest frequencies) are considered. Since usually $r \ll n$, we have a reduced-order model.

The *reduced-order reference model* (or simply *reduced-order model* - ROM) can be assembled by substituting the generalized coordinates with the modal coordinates in the full-order model given by Equation 7, considering the variable transformation $\mathbf{q} = \boldsymbol{\Phi}_r \boldsymbol{\eta}$ and yielding:

$$\dot{\boldsymbol{\eta}} = \left[\boldsymbol{\Phi}_r^{\top}\left(\mathbf{J}^{\top}\,\mathbf{M}\,\mathbf{J} + \mathbf{I}\right)\boldsymbol{\Phi}_r\right]^{-1}\mathbf{p}_r,$$
$$\dot{\mathbf{p}}_r = -\,\boldsymbol{\Phi}_r^{\top}\mathbf{K}\boldsymbol{\Phi}_r\,\boldsymbol{\eta} - \boldsymbol{\Phi}_r^{\top}\mathbf{C}\boldsymbol{\Phi}_r\,\dot{\boldsymbol{\eta}} + \boldsymbol{\Phi}_r^{\top}\mathbf{Q}, \tag{10}$$

where $\mathbf{p}_r$ is referred to as modal momenta and is consistent with the variable transformation applied to the generalized coordinates. Moreover, the Jacobian matrix in this case is a function of the modal coordinates, $\mathbf{J} = f(\boldsymbol{\eta})$.

3 HAMILTONIAN NEURAL NETWORKS

Conventional feed-forward neural networks, also known as multi-layer perceptrons (MLPs), learn and generalize directly from data, not relying on any physical constraint. For a Hamiltonian system, a well-trained MLP model can predict the time derivatives of the generalized coordinates and momenta, as depicted in Figure 5.

MLP models, however, tend to cause energy to drift with time, not adhering to the energy conservation principle. A solution to this problem has been proposed by [8]: to parameterize the Hamiltonian function with a neural network, learning it directly from data, instead of crafting it by hand using domain-specific knowledge. This architecture was called the Hamiltonian neural network (HNN).

In their forward pass, Hamiltonian neural networks receive generalized coordinates and momenta as inputs and, using an MLP as a subnetwork, parameterize an energy-like scalar $\hat{\mathcal{H}}$. The gradients of this scalar are taken with respect to the inputs using automatic differentiation, as illustrated in Figure 6. Finally, we compute and optimize the following L_2 loss function:
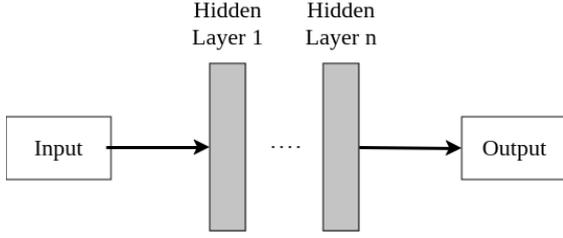

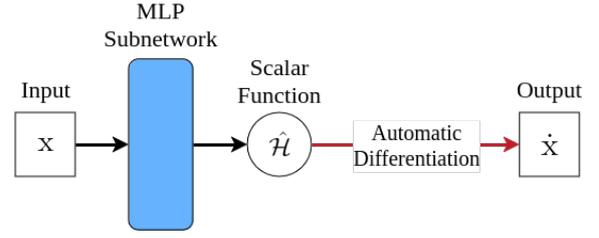Figure 5: Multi-layer perceptron (MLP) architecture.

Figure 6: Hamiltonian neural network architecture.

For HNN models, the following L_2 loss function has to be minimized:

$$\mathcal{L} = \left\| \frac{\partial \hat{\mathcal{H}}}{\partial \mathbf{p}} - \frac{\partial \mathbf{q}}{\partial t} \right\|_2 + \left\| \frac{\partial \hat{\mathcal{H}}}{\partial \mathbf{q}} + \frac{\partial \mathbf{p}}{\partial t} \right\|_2. \tag{11}$$

With this strategy, HNNs learn conserved quantities analogous to the total energy directly from data and in an unsupervised manner. Ref. [8] has found other interesting properties of HNNs besides the conservation laws: HNNs are perfectly reversible in time - the mapping of the generalized variables at one time to another time is bijective - and the conserved quantity (analogous to the total energy) can be manipulated by integrating it along the gradient of the Hamiltonian.

A slightly different neural network architecture, named dissipative-Hamiltonian neural network (D-HNN), was proposed in [9]. The idea is to extend the HNNs by including a second MLP subnetwork to separately learn a scalar $\hat{\mathcal{D}}$, analogous to a dissipative function, as depicted in Figure 7. In the D-HNN framework, an implicit Helmholtz decomposition of the vector field is performed, separately extracting its conservative (rotational) and dissipative (irrotational) components.
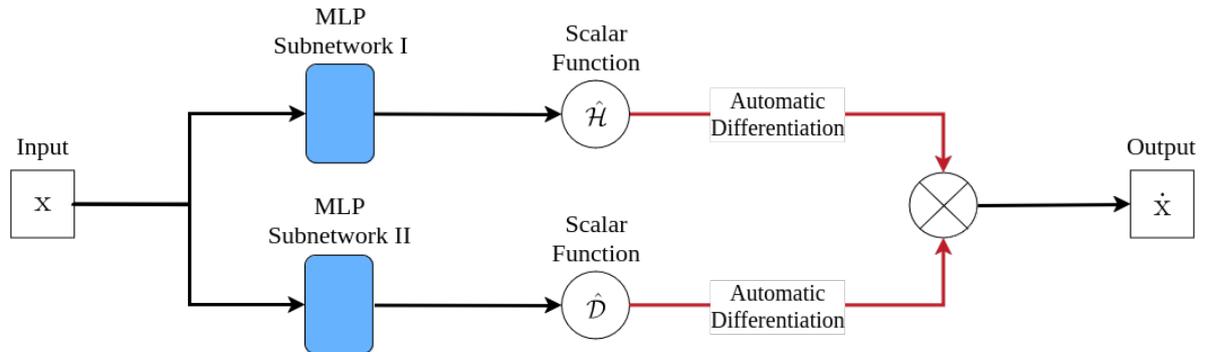

Figure 7: Dissipative-Hamiltonian neural network architecture.

6

For D-HNNs, the L_2 loss function in this case assumes the following form:

$$\mathcal{L} = \left\| \left(\frac{\partial \hat{\mathcal{H}}}{\partial \mathbf{p}} + \frac{\partial \hat{\mathcal{D}}}{\partial \mathbf{q}} \right) - \frac{\partial \mathbf{q}}{\partial t} \right\|_2 + \left\| \left(-\frac{\partial \hat{\mathcal{H}}}{\partial \mathbf{q}} + \frac{\partial \hat{\mathcal{D}}}{\partial \mathbf{p}} \right) - \frac{\partial \mathbf{p}}{\partial t} \right\|_2 . \tag{12}$$

The minimization of the loss function for HNN and D-HNNs follows the same procedure as for an MLP model and can be achieved by using an optimization algorithm, such as stochastic gradient descent.

4 RESULTS

In this section, the main findings of this study are presented. First, we show validation results for the reference models derived in Section 2. Next, we describe the datasets and training hyperparameters chosen for the neural network models. And, finally, we explore the performance of the surrogate models based on Hamiltonian neural networks when applied to highly flexible beams. We simulate free and forced responses for a highly flexible cantilever beam and a highly flexible beam-like wing, comparing the computational cost (wall-clock simulation time) and accuracy relative to reference and baseline models.

4.1 Validation of Reference Models

We investigate how well the full- and reduced-order reference models derived in Section 2 can represent the dynamics of a highly flexible cantilever beam considered in [15, 16]. The relevant properties of the structure are summarized in Table 1. We simulated the out-of-plane bending

Table 1: Properties of the highly flexible cantilever beam.

Linear mass density	0.2 kg/m
Length	1 m
Out-of-plane bending stiffness	50 N·m^2

response of the structure acted upon the same load conditions investigated by [16], assuming their results as our ground truth. Unfortunately, since the actual numerical data was unavailable, we estimated it from the graphs presented in their work. Hence, only a qualitative analysis was possible.

Setting a full-order model with $N = 10$ and a reduced-order model with $r = 1$, we created a simulation framework in Python using an implicit Runge-Kutta integration method of the Radau IIA family of 5th-order (*Radau* method of the *SciPy* library) considering default tolerances.

For the static analysis, we compared the tip position under different upward load magnitudes F_z acting at the free tip of the beam. In the dynamic analysis, we compared the time history of the tip position considering a time-dependent vertical tip load $F_z = 10 \sin (50t)$, intentionally exciting the *beat effect* over the structure. Figures 8 and 9 depict validation examples in which both numerical models succeeded in representing static and dynamic behaviors of the structure, respectively.

The shown examples showcase the ability of the proposed models to accurately represent geometrical nonlinearities, which can be seen by out-of-plane deflections (z-axis) equivalent to more than 50% of the length of the beam.
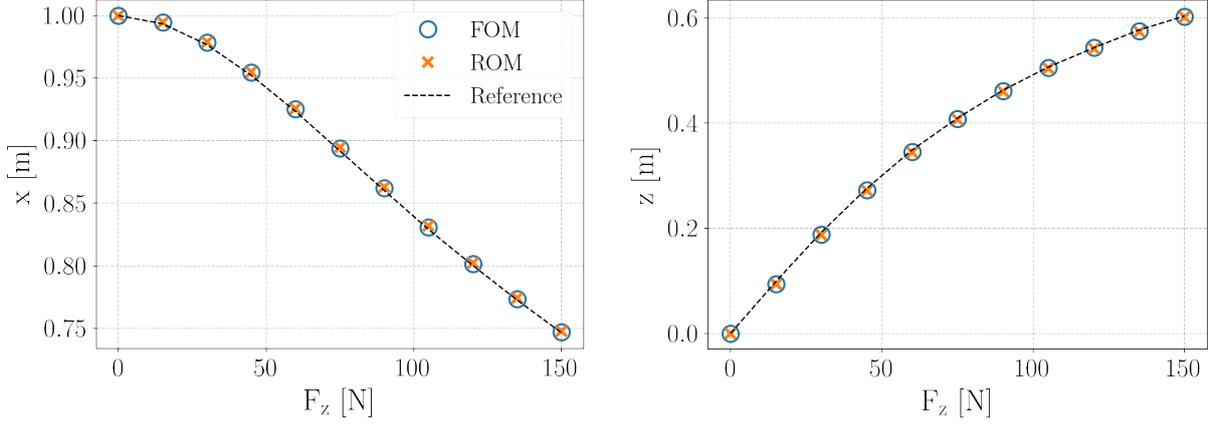
Figure 8: Equilibrium tip position under different static vertical tip load conditions.
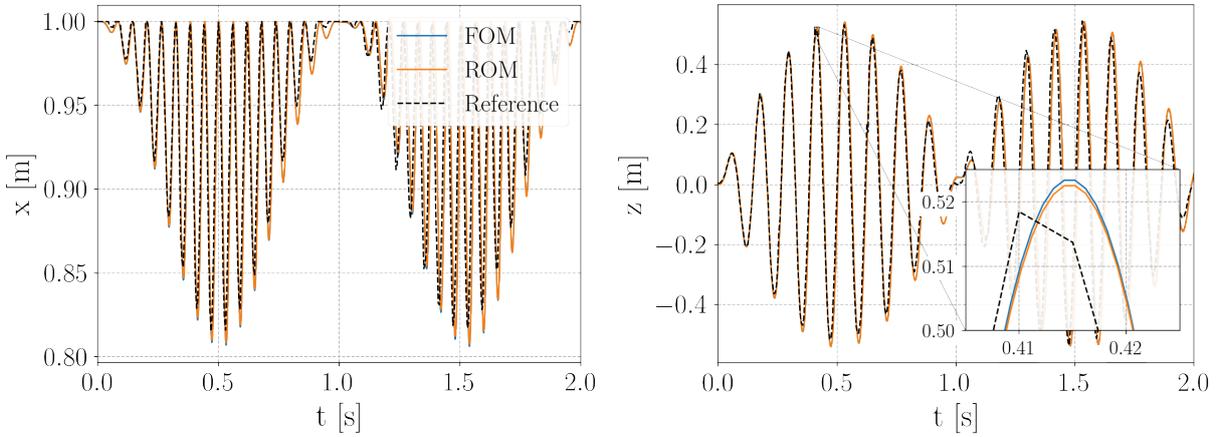

Figure 9: Tip position time history for the structure under the dynamic load condition.

The second structure we consider in this study is the highly flexible beam-like wing proposed by [17] for aeroelastic analyses. Unlike the original structure, however, we modified some properties and included an offset between the elastic axis and center of gravity, allowing torsion deformations to develop. The main properties of the highly flexible wing are summarized in Table 2.

Table 2: Properties of the highly flexible wing.

Linear density	0.75 kg/m
Length	16 m
Chord (c)	1 m
Mass moment of inertia	0.75 kg·m^2
Chordwise elastic-axis	$0.5c$ m
Chordwise center of gravity	$0.267c$ m
Out-of-plane bending stiffness	2×10^4 N·m^2
Torsion stiffness	1×10^3 N·m^2

To validate the full- and reduced-order models derived for the wing, we compare the frequencies of the first 9 modes of vibration considering the results obtained by [12] as reference. A FOM with $n = 16$ and a ROM with $r = 3$ are assumed in this analysis.

Table 3: Comparison of natural frequencies (in rad/s) of the system using different models.

Mode	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
FOM	2.21	9.55	13.78	27.39	36.38	46.69	63.14	65.05	80.53
ROM	2.21	9.59	13.83	27.63	37.02	47.14	64.28	67.54	84.65
Reference	2.21	9.55	13.78	27.39	36.38	46.69	63.14	65.05	80.53

The results shown in Table 3 emphasize the capability of both models in representing the dynamics of the wing when bending and torsion are coupled, especially for the reduced-order model. However, it can be noted that as frequencies increase, the ROM solution diverges from the reference values. The full-order modeling technique used in this study is the same one that produced the reference values in [12], thus the matching values seen in the table were predictable.

4.2 Datasets and Training

The neural network models described in Section 3, namely MLP, HNN, and D-HNN, were trained, generating the respective surrogate models for the systems we assumed as subjects of our study. The artificial datasets used for training the neural networks were based on simulations of the validated reduced-order models. While we considered a ROM with $r = 1$ for the cantilever beam (obtained from a FOM with 10 elements), $r = 3$ was assumed for the highly flexible wing (obtained from a FOM with 16 elements), analogously to the previous subsection.

The dataset for the highly flexible beam consists of the generalized coordinates and momenta as input features and their time derivatives as outputs, which were obtained from 50 free-response simulations considering pseudo-random initial conditions such that $-\pi/9 \leq q_k \leq \pi/9$ rad and $0 \leq p_k \leq 1.5$ kg·m^2/s. The simulation framework described in the previous subsection was employed to run the simulations for $0 \leq t \leq 0.2$ s using a timestep of 0.001 s. We used 80% of the samples for training and the remaining for validation. Moreover, since the cantilever beam is not subjected to structural damping, the D-HNN model would unnecessarily try to estimate a dissipation function, therefore we did not consider it when analyzing this structure.

We aimed to compare the differences in performance of the different neural networks not only relative to the reference model but also to the baseline MLP model, hence we decided all neural network models would be set with the same hyperparameters. The hyperparameters were chosen after testing different combinations of the number of neurons, activation functions, batch size, and so on, in a tedious and perhaps the most challenging procedure of this study. We defined the MLP and HNN models as having a single fully-connected layer with 128 neurons and ELU activation function, which were trained for $5,000$ epochs in batches of size 64 using the *Adam* algorithm with an initial learning rate of 1×10^{-5}. At the end of the training, both training and validation losses converged such that $\mathcal{L} \propto 10^{-7}$.

Analogously to the procedure for the cantilever beam, the dataset for the highly flexible wing was created but considering simulations for $0 \leq t \leq 1$ s. We assumed structural damping in the form described in Equation (5), such that $\beta_k = 0.1$ and $\beta_k = 10$ for the out-of-plane and torsion degrees-of-freedom. The HNN model would not be adequate for this analysis since it does not account for dissipation effects, therefore we only considered the baseline MLP and the D-HNN models. After repeating the procedure for choosing the hyperparameters mentioned previously, we defined both models as having a single fully-connected layer with 256 neurons and ELU activation function, which were trained for $20,000$ epochs in batches of size 64 using

the *Adam* algorithm with an initial learning rate of 1×10^{-4}. This time, training and validation losses converged such that $\mathcal{L} \propto 10^{-6}$.

4.3 Surrogate Models Using Hamiltonian Neural Networks

The surrogate models obtained by training the different neural networks mentioned in the last subsection were used to simulate the two structures considered under different free and forced response conditions.

We simulated the free response of the highly flexible cantilever beam for 30 seconds considering 10 different pseudo-random initial conditions. The mean and standard deviation of the computational cost (wall-clock time) relative to the reference model (ROM) was $1.29\% \pm 0.07\%$ for the HNN model and $0.78\% \pm 0.06\%$ for the baseline MLP model. The simpler architecture of the MLP certainly makes it faster to infer, explaining its better computational performance. Figure 10 shows the first 2 s of a simulation considering an arbitrary initial condition. The accuracy achieved by both surrogate models is remarkable, at least for short-term simulation.
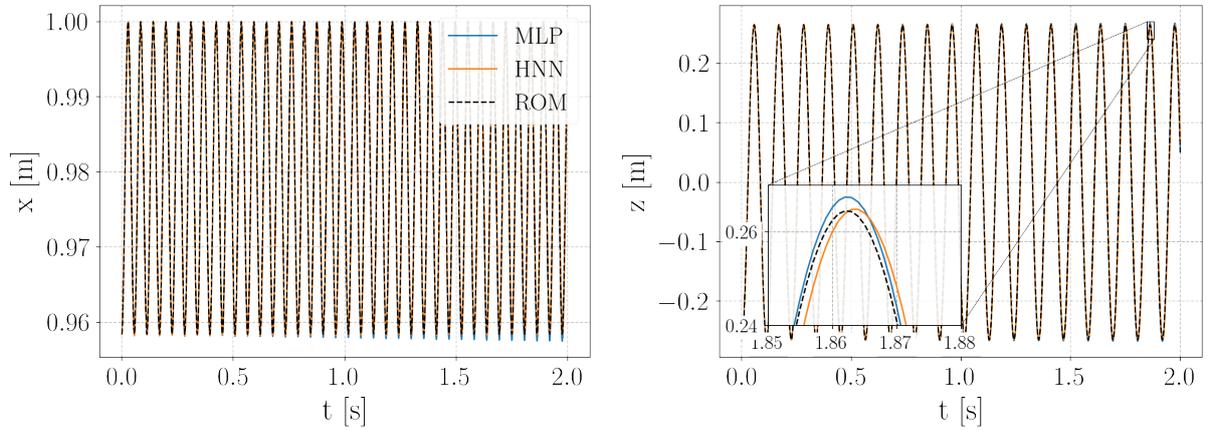

Figure 10: Tip position of the structure for an arbitrary free response simulation.

The neural network-based models were trained on datasets generated from free response simulations. Therefore, if we aim to use the surrogate models for cases involving external loads applied to the system, an additional function call is required to calculate the corresponding generalized forces analytically and add them to the predicted time derivatives of the generalized momenta. Figure 11 illustrates the modified structure of this surrogate model, which we refer to as the augmented surrogate model.
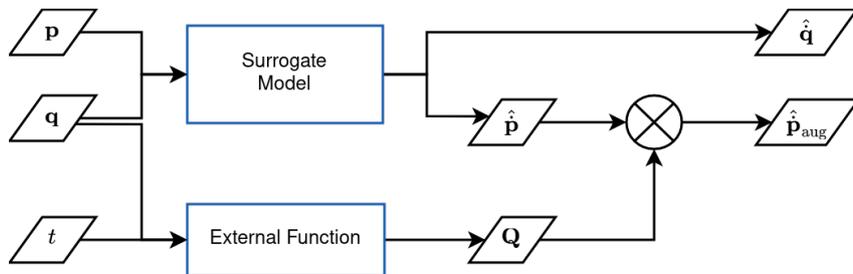

Figure 11: Augmented surrogate model to account for external loads.

Using the augmented structure, we conducted 10 forced response simulations of 30 s considering the same dynamic load condition used during the validation of the reference model, however,

the load was applied for an arbitrary period such that $t \leq 5$ s.

The mean and standard deviation of the computational costs relative to the reference model for the forced response simulations were $83.60\% \pm 4.63\%$ for the HNN models and $82.72\% \pm 4.34\%$ for the MLP model. The surrogate models are still more efficient than the reference model, however, the performance gains were much lower due to the need for the external function calls to compute the generalized forces term.

Figure 12 depicts the simulation of the structure acted by the dynamic load condition for 2 s, analogously to the conditions considered to produce the results during the validation step.
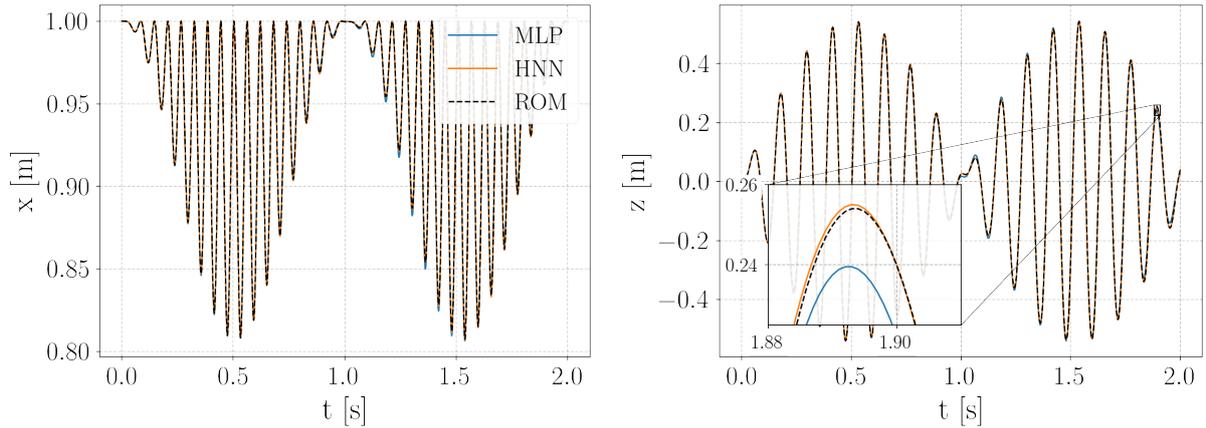
Figure 12: Tip position time history for the structure under a dynamic load condition.

The solution produced by the MLP seems to be not as accurate as it was for the free response. We investigated this behavior and found that for all simulations tested the MLP tends to diverge if we consider longer simulations. Unlike the HNN model, the MLP model does not preserve the total energy of the system for long simulations, as depicted in Figure 13 for a simulation considering the dynamic load acting for an arbitrary period. On the other hand, the HNN solution fluctuates (possibly caused by the integration error) around the ground truth for the whole simulation.
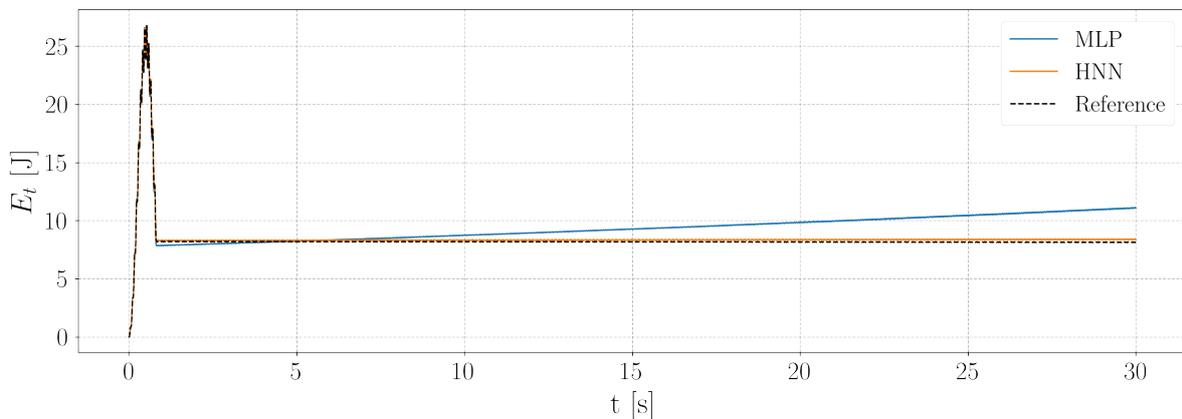

Figure 13: Total energy of the system as a function of time.

Turning our attention to the highly flexible wing structure, we used the MLP and D-HNN surrogate models to simulate the free response of the structure also for 10 different pseudo-random

initial conditions. An arbitrary solution is shown in Figure 14 in the form of time histories of the tip position and twist angle.
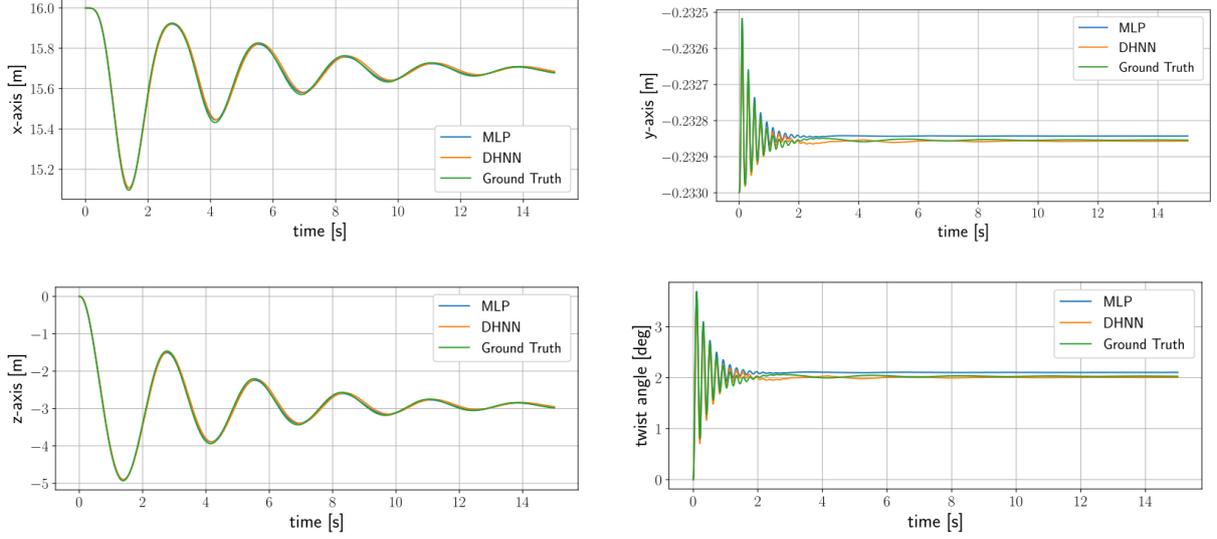

Figure 14: Tip position and twist angle time histories for an arbitrary initial condition.

The relative computational cost of the D-HNN model was $12.38\% \pm 4.07\%$, while for the MLP model, it was $8.99\% \pm 4.13\%$, both relative to the reference model. Therefore, the computational costs associated with the MLP are slightly advantageous for the test batch considered.

Considering the augmented surrogate model architecture presented previously, we conducted 10 simulations considering the application of the following tip force components:

$$F_{i,y} = \int_0^L 0.5t^2 \left[1 - \sqrt{1 - \left(\frac{x}{L}\right)^2} \right] dx, \tag{13}$$

$$F_{i,z} = \int_0^L 2t^2 \sqrt{1 - \left(\frac{x}{L}\right)^2} \, dx. \tag{14}$$

The proposed forces mimic the lift and drag distributions over the wing and are applied for a random period such that $t \leq 5$. The tip position and twist angle time histories considering the forces being applied for 5 seconds are depicted in Figure 15.

Both the MLP and D-HNN surrogate models again demonstrate similar accuracy. In this case, the performance gains were considerably lower than for the free response case due to the need for the analytical computation of the generalized forces. The relative computational cost of the D-HNN model was $51.91\% \pm 5.84\%$, while it was $56.39\% \pm 2.56\%$ for the MLP model.

Because of the structural damping, the energy drifting observed during the cantilever beam analysis could not be seen for the highly flexible wing case. In such cases, the simpler MLP architecture is faster than the D-HNN with comparable accuracy. However, it is worth highlighting that the Hamiltonian neural networks can be powerful tools for modeling dynamical systems when energy conservation is a requisite. They provide a level of interpretability that is not achievable with simple feed-forward models, making them more reliable options for engineering applications.
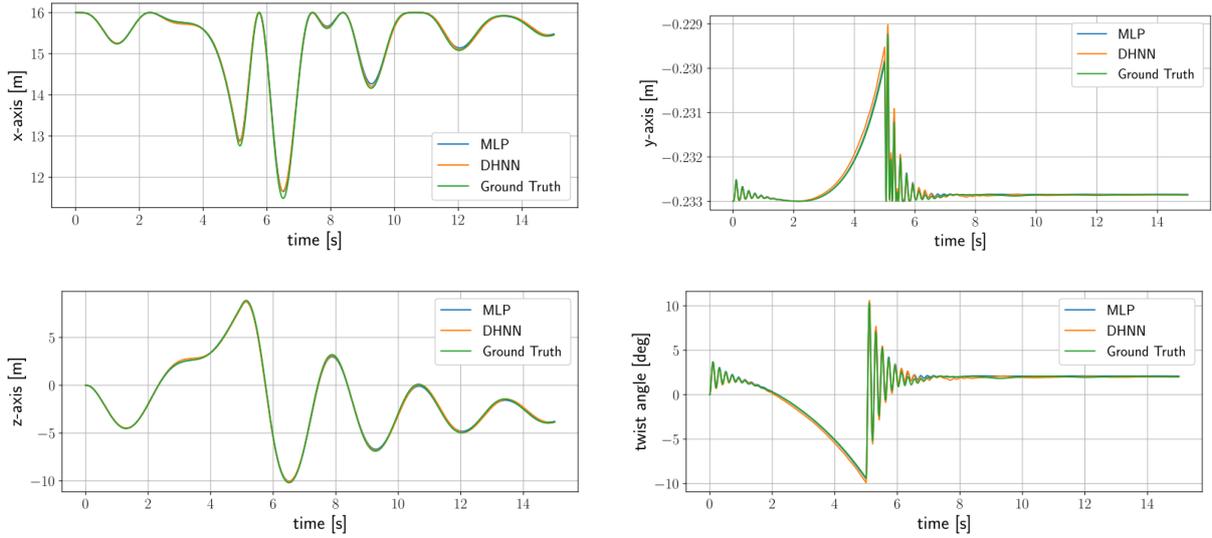
Figure 15: Tip position and twist angle time histories for the structure subjected to a dynamic tip load.

In this study, we explored only systems with a small number of degrees of freedom. However, we noted that when considering more complex systems, choosing the hyperparameters for the neural networks becomes a challenging task, and the training times increase exponentially. Therefore, further investigations are necessary to determine the applicability of the proposed methodology for systems more complex than those we considered.

5 USING NEURAL NETWORKS FOR SURROGATE MODELING THE VORTEX-LATTICE METHOD

The intricate coupling between aerodynamic forces and structural deformations present in the dynamics of highly flexible aircraft demand modeling techniques of the highest complexity, where simplifications are often inappropriate. Regarding aerodynamics modeling, unsteady models that consider high angles of attack should be preferred [3], which often leads to prohibitive computational costs.

Even for (non-highly) flexible aircraft, only a few methods can produce faster than real-time simulations. An example can be found in [18] via combined algebraic optimizations of incremental aerodynamic forces and linear reduced-order models for the aeroelastic and aerodynamic lag state equations.

Preliminary findings from [19] indicate that the aerodynamics subroutine based on the vortex-lattice method (VLM) is responsible for roughly 99% of the total computational cost to simulate a flexible aircraft. In this analysis, however, the (linear) Rayleigh-Ritz method was chosen for the dynamic-structural model, hence mass, stiffness, and aeroelastic transformation matrices are calculated before the simulation. Despite that, this work introduced how simple feed-forward neural networks can be used for the aerodynamics surrogate modeling of a fairly complex flexible aircraft developed originally by [20] and called Generic Narrow-Body Airliner (GNBA).

The GNBA has 32.756 m of wing span, a fuselage length of 39.15 m, and a total mass of $55,764$ kg in the design weight (DW) configuration, which includes the operational empty weight (OEW), payload, and fuel. A 3-dimensional representation of the GNBA and its DW mass distribution are depicted in Figures 16 and 17, respectively.
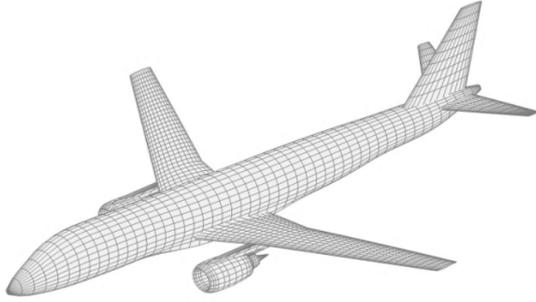
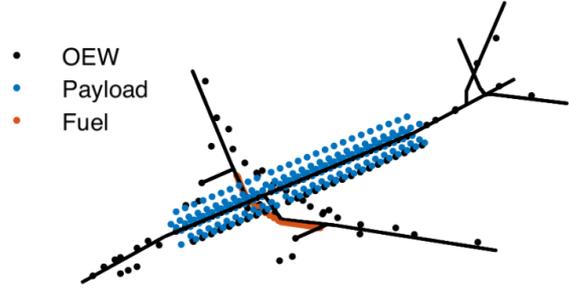Figure 16: The Generic Narrow-Body Airliner [20].

Figure 17: GNBA design weight configuration [20].

In this section, we employ most of the methodology described in [19] regarding the use of neural networks to surrogate modeling the VLM for the GNBA. Two differences can be noted in the present approach: (i) instead of using four different neural networks (one for each output group: drag, longitudinal, lateral-directional, and generalized modal aerodynamic forces), we concluded only one neural network is necessary; and (ii) instead of using the elegant Sobol sequence for sampling, we simply assumed a pseudo-random combination of all features to sample our input data.

Flexibility effects are included in the solution of the semi-steady VLM model by considering structural displacements and velocities represented via modal coordinates $\boldsymbol{\eta}$ and velocities $\dot{\boldsymbol{\eta}}$ obtained from a reduced-order model assuming the six most influential modes of vibration on the aerodynamic forces, namely the first, second, and fifth to eighth modes. As investigated in [19], the third and fourth modes have little influence over the aerodynamic forces, allowing us to neglect them.

The input data sum up to 25 features: Mach number M, altitude h, angle of attack α, angle of sideslip β, reduced roll rate p/V, reduced pitch rate q/V, reduced yaw rate r/V (where V is the aerodynamic speed), horizontal tail incidence angle i_t, left elevator deflection angle $\delta_{e,L}$, right elevator deflection angle $\delta_{e,R}$, left aileron deflection angle $\delta_{a,L}$, right aileron deflection angle $\delta_{a,R}$, rudder deflection angle δ_r, the six modal coordinates $\boldsymbol{\eta}$, and the six reduced modal velocities $\dot{\boldsymbol{\eta}}/V$. Such variables were chosen due to their relevance in the VLM code for outputting the following 12 aerodynamic coefficients: lift C_L, drag C_D, side force C_Y, pitching moment C_m, rolling moment C_l, yawing moment C_n, and the six generalized modal aerodynamic forces $\mathbf{C}_\varphi$. The limits considered for each feature and the reasoning behind them can be found in [19].

Our dataset contains $4,096$ samples split in 80% for training and the remaining for validation. The neural network architecture consists of a single hidden layer with 75 neurons and hyperbolic tangent as the activation function. The training was carried out for 250 epochs by the *Levenberg-Marquardt* algorithm on the Matlab® Deep Learning Toolbox™.

After training, the surrogate model is included in the simulation framework, replacing the VLM subroutine. Considering the equilibrium state at straight and level flight at $M = 0.7$ and $h = 10,000$ m, we linearized the equations of motion and obtained the eigenvalues corresponding to the dynamical modes of the system, as depicted in Figure 18, comparing the results with those from the VLM.

The proposed neural network predicted the dynamical modes with remarkable accuracy, in which the major numerical discrepancy is seen for the 8th aeroelastic mode. The spiral mode
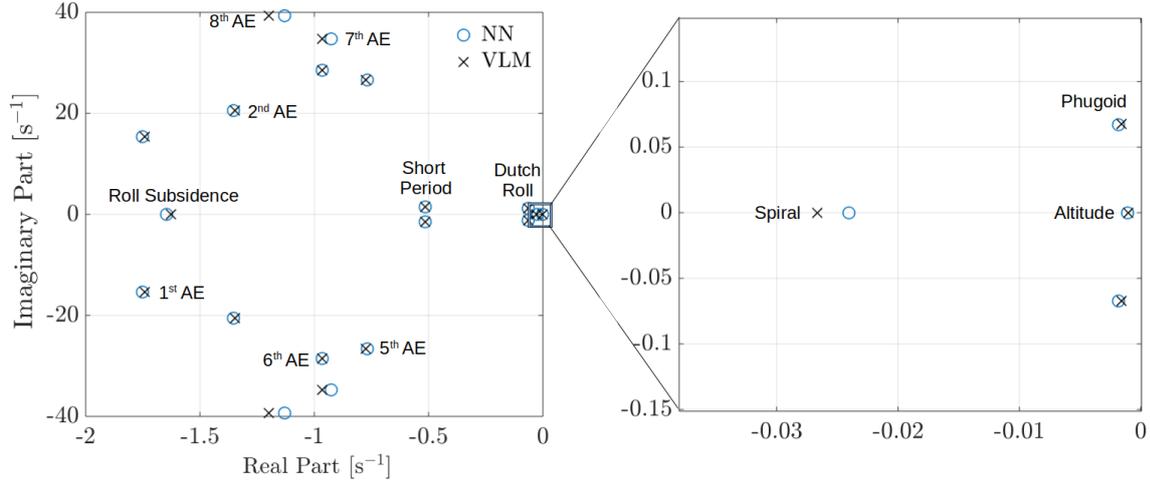
Figure 18: Eingevalues of classical (rigid body) and aeroelastic modes around the trimmed condition.

also shows a reasonable distance from its original value, however, it is influenced by the scale chosen for the graph, since numerically the difference is not expressive. A more in-depth study would be necessary to investigate how the reported differences might impact, for example, the design of control laws.

Further assessment of the surrogate model was made in a series of simulations of the nonlinear equations of motion for different command inputs. We used a 4th-order Runge-Kutta integration scheme in a Matlab® simulation framework, which was slightly adapted from the work of [19]. Figures 19-21 show the responses of longitudinal, lateral-directional, and aeroelastic states to an elevator doublet command in which $\Delta\delta_{e,R} = \Delta\delta_{e,L} = 5°$ for $0.5 < t \leq 1.5$ s and $\Delta\delta_{e,R} = \Delta\delta_{e,L} = -5°$ for $1.5 < t \leq 2.5$ s.

The longitudinal states represented in Figure 19 are x-axis inertial velocity component u, z-axis inertial velocity component w (both in the body-fixed reference frame), pitch rate q, pitch angle θ, altitude h, and x-axis position x (in the inertial frame). The response results show a very close relation between the neural network-based and the original values of the longitudinal states for the period simulated.

On the other hand, the differences between the two approaches are visible for the lateral-directional states, namely y-axis inertial velocity component v, roll angle ϕ, roll rate p, yaw rate r, yaw angle ψ, and y-axis position y (in the inertial frame). Nonetheless, one may note the variations of the states shown in Figure 20 are orders of magnitude smaller than those observed for the longitudinal states. Therefore, the differences pointed out are of small relevance for short simulations in this particular scenario, once the elevator doublet mainly excites the longitudinal states.

Finally, Figure 21 depicts the response of the modal displacements to the elevator command. Once again, the significant differences are of small magnitude and observed for the 2nd, 5th, and 8th aeroelastic modes, which are asymmetric and, therefore more sensitive to lateral-directional excitation. The symmetrical modes (1st, 6th, and 7th) have very similar responses for both methods, showcasing the good accuracy levels of the solutions produced by the neural network-based model. It must be noted, however, that the differences observed may lead to compromising errors if long simulations are performed, such that further investigation is advised.
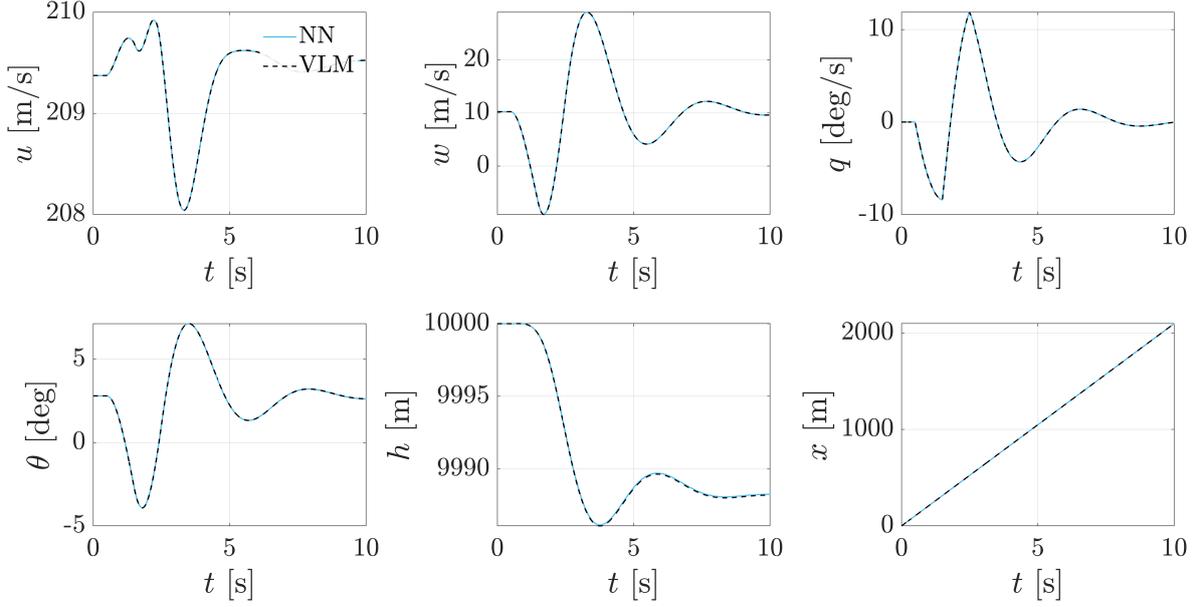
Figure 19: Longitudinal states response due to the elevator doublet command.
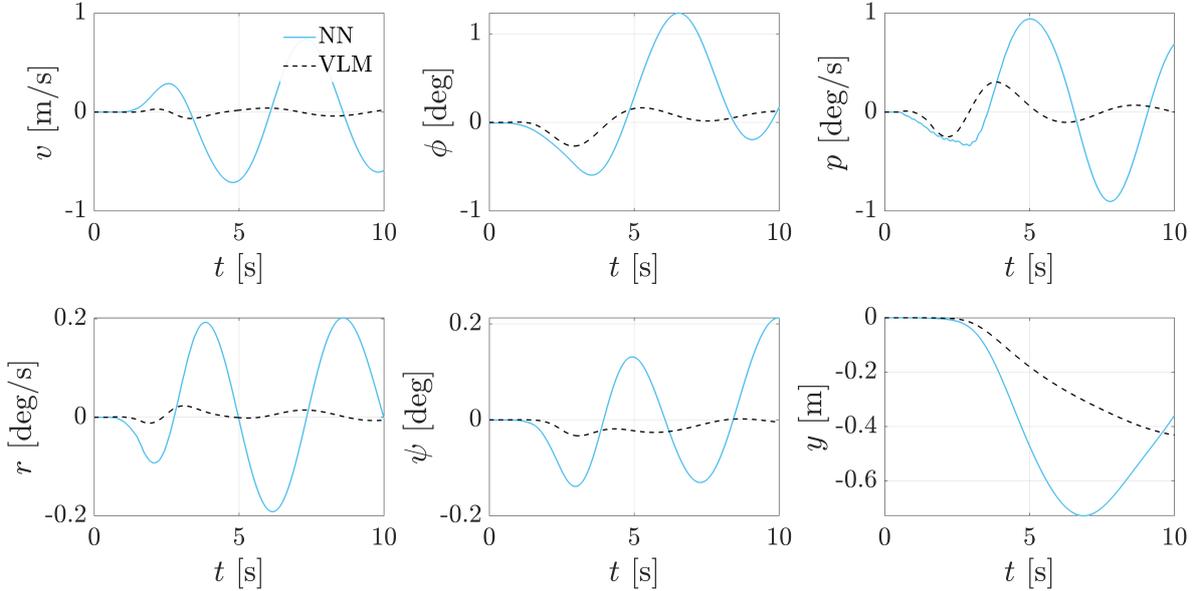

Figure 20: Lateral-Directional states response due to the elevator doublet command.

The computational costs of simulating the GNBA were investigated by analyzing its response to elevator, aileron, and rudder doublet commands for random trimmed states and command deflection angles (within the limits used for the dataset) applied for the same time interval mentioned previously. Using an Intel® Core™ i7-11800H @ 2.30GHz×16 CPU and the same simulation framework described previously, we conducted 90 simulations (30 for each command type), each lasting 5 s with a time step of 0.005 s.

During preliminary analyses, we found that the surrogate aerodynamic models could achieve a reduction in simulation time of approximately 45%, which was lower than we expected. To accelerate the neural network inference, we converted the neural network into a regular Matlab function using the *genFunction* function. By doing this, we could finally see the expected performance improvement at the cost of some accuracy loss. After some tests, we found that the accuracy loss is proportional to 10^{-12}, therefore negligible.
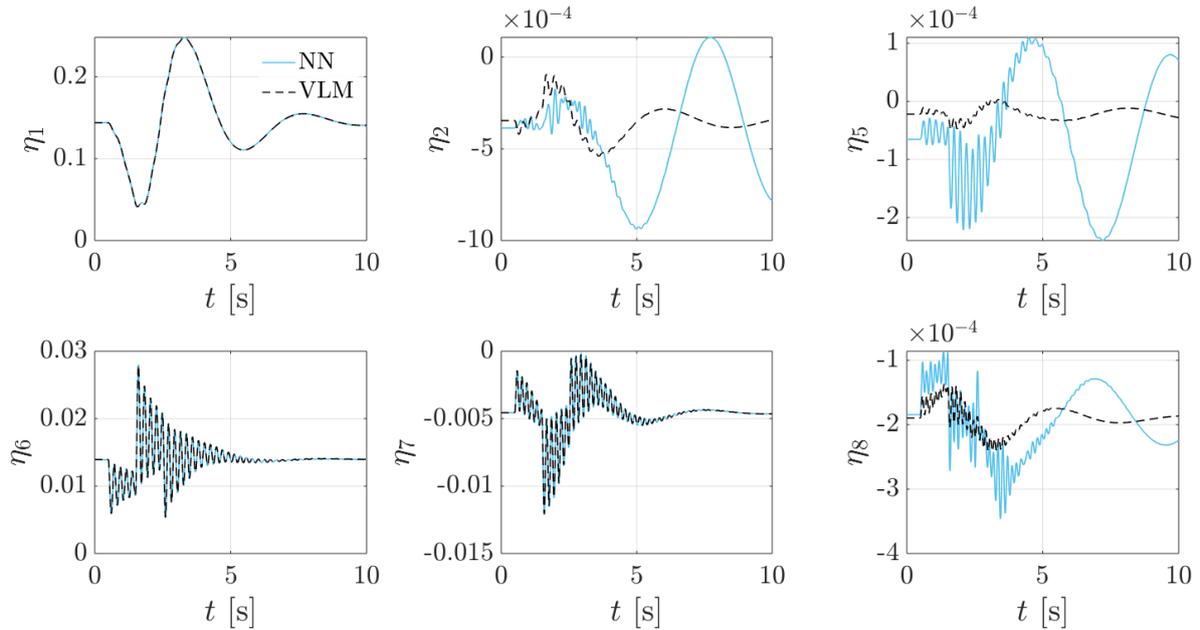
Figure 21: Aeroelastic states response due to the elevator doublet command.

After running the 90 proposed simulations, the relative computational cost (ratio of wall-clock time) between the surrogate model and the original VLM was $0.995\% \pm 0.025\%$. This means an increase in computational performance by a factor of 100, on average. For instance, it can take as little as 0.56 s in our CPU to simulate 5 s of the aircraft response by using the surrogate aerodynamic model, instead of 56.7 s by using the VLM. This remarkable achievement opens up a series of perspectives on the use of neural network-based models where real-time simulations are necessary.

The work of [19] extends to more elaborate uses of neural networks by considering the linearization of the VLM with respect to modal coordinates and the inclusion of second-order terms present in the equations of motion as features of the neural network model, which proved to be beneficial when the flexibility of the structure is increased.

Future research perspectives in this area are aimed at taking a step further and investigating how arbitrarily (highly) flexible aircraft can leverage the benefits of surrogate modeling based on neural networks. A promising area of investigation is the combination of dynamic-structural surrogate models suitable for geometrically nonlinear structures, such as the one showcased in Section 4, to more intricate aerodynamic models, for instance, the double-lattice method (DLM) and stall models, which may be modeled using physics-informed neural networks (PINNs).

6 CONCLUSIONS AND FURTHER WORK

The increasing concerns about climate change are fueling disruptive technologies in the aeronautical industry. Higher aspect ratio lifting surfaces can lead to significant improvements in aerodynamic performance, however, such structures have become more and more flexible, posing challenges for traditional modeling techniques and requiring the use of computationally demanding solutions. To reduce the computational costs associated with the simulation of highly flexible structures, we proposed to use Hamiltonian neural networks (HNNs) as surrogate models, leveraging the energy-preserving character of the Hamiltonian formalism into the framework of neural networks. We modeled two highly flexible beam-like structures using a lumped-mass multibody approach and derived a reduced-order model based on modal

decomposition while preserving the geometrical nonlinearities with the use of exact kinematic relations. After validating the reduced-order models, we assumed them as the reference models used to create the dataset and train the neural network models. We compared the surrogate models in terms of efficiency and accuracy against the reference model and a baseline surrogate model based on a feed-forward (multi-layer perception - MLP) neural network.

The results indicate that the HNN surrogate models can drastically reduce the computational costs associated with free response simulations: up to approximately 98% for the highly flexible beam structure and up to 87% for the highly flexible beam structure. Even though the MLP models can provide an additional $1-6\%$ reduction in the computational cost, they are not able to preserve energy, being unreliable for long-term simulations. On the other hand, a well-trained HNN learns the energy conservation laws directly from data and delivers solutions of the same order of accuracy as the reference model, even for long simulations. For forced-response simulations, the architectures used for the surrogate models demand the computation of the generalized forces in an external function, which ends up compromising the efficiency gains obtained by the proposed approach.

We also investigated how neural networks can be used to surrogate modeling the aerodynamics of flexible aircraft, particularly substituting the vortex-lattice method (VLM). Using a flexible aircraft simulation framework available, we trained a neural network to predict the aerodynamic coefficients, including the generalized modal aerodynamic forces caused by structural flexibility. Employing the neural network model we simulated the aircraft under 90 different trim conditions and command inputs. The results show that not only did the neural networks provide reasonably accurate solutions but can also improve the simulation framework efficiency by an impressive factor of 100, on average.

In this study, we presented some interesting perspectives on the use of neural networks in the context of flexible and highly flexible aircraft. We hope these perspectives enable deeper advancements in the area. Some topics for future investigation include removing the necessity of analytically computing the generalized forces for the surrogate models, the use of data-free supervised learning models, and finally, the combination of dynamic-structural and aerodynamic surrogate models based on neural networks. Given the current state of the art, physics-informed neural networks seem good candidates to produce interpretable and generalizable surrogate models.

ACKNOWLEDGMENTS

7 REFERENCES

[1] Kilimtzidis, S. and Kostopoulos, V. (2023). Multidisciplinary structural optimization of novel high-aspect ratio composite aircraft wings. *Structural and Multidisciplinary Optimization*, 7(150).

[2] Laboratory of New Concepts in Aeronautics (2019). X-HALE-BR prototype. Available at: https://sites.google.com/view/labnewconcepts.

[3] Hodges, D., Chang, C. S., and Patil, M. (2008). Flight dynamics of highly flexible aircraft. *Journal of Aircraft*, 45, 538–545.

[4] Shearer, C. M. and Cesnik, C. E. S. (2007). Nonlinear flight dynamics of very flexible aircraft. *Journal of Aircraft*, 44(5), 1528–1545.

[5] Guimarães Neto, A. B., Cardoso-Ribeiro, F. L., and Silvestre, F. J. (2018). Applicability of geometrically-linear structural-dynamic models for the flight dynamics of arbitrarily-flexible aircraft. In *Proceedings of the 31st Congress of the International Council of the Aeronautical Sciences, ICAS 2018*. p. 14.

[6] Baker, N., Alexander, F., Bremer, T., et al. (2019). Workshop report on basic research needs for scientific machine learning: core technologies for artificial intelligence. Report, U.S. Department of Energy, ASCR.

[7] Cranmer, M., Greydanus, S., Hoyer, S., et al. (2020). Lagrangian neural networks. *Computing Research Repository - CoRR*.

[8] Greydanus, S., Dzamba, M., and Yosinski, J. (2019). Hamiltonian neural networks. *Computing Research Repository - CoRR*.

[9] Sosanya, A. and Greydanus, S. (2022). Dissipative hamiltonian neural networks: learning dissipative and conservative dynamics separately. *Computing Research Repository - CoRR*.

[10] Mattheakis, M., Sondak, D., Dogra, A. S., et al. (2022). Hamiltonian neural networks for solving equations of motion. *Physical Review E*, 105(6). doi:10.1103/physreve.105.065305.

[11] Desai, S. A., Mattheakis, M., Sondak, D., et al. (2021). Port-hamiltonian neural networks for learning explicit time-dependent dynamical systems. *Physical Review E*, 104.

[12] Veronese, A. L. (2021). *Structural model order reduction for flexible aircraft simulation*. Master of science in aeronautical and mechanical engineering, Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, Brazil.

[13] Rempel, M., Cardoso-Ribeiro, F. L., and Moreira, F. (2020). Lumped element multi-body modeling approach for very flexible aircraft. In *Proceedings of the XLI Ibero-Latin-American Congress on Computational Methods in Engineering*. Foz do Iguaçu, Paraná, Brazil, p. 7.

[14] Squire, W. and Trapp, G. (1998). Using complex variables to estimate derivatives of real functions. *SIAM Review*, 40(1), 110–112.

[15] Cheng, T. (2002). *Structural Dynamics Modeling of Helicopter Blades for Computational Aeroelasticity*. Master of science in aeronautics and astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

[16] Brown, E. L. (2003). *Integrated Strain Actuation in Aircraft with Highly Flexible Composite Wings*. Doctor of science in mechanical engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

[17] Patil, M. J. (1999). *Nonlinear Aeroelastic Analysis, Flight Dynamics, and Control of a Complete Aircraft*. Doctor of philosophy in aerospace engineering, Georgia Institute of Technology, Atlanta, Georgia, USA.

[18] Paulino, J. A. (2020). *Studies on Model Order Reduction for Fast Simulation of Flexible Aircraft*. Doctor of science in aeronautical and mechanical engineering, Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, Brazil.

[19] Vieira, B. S. C. (2021). *Flight Simulation of a Flexible Aircraft Using Neural Networks for Real-Time Applications*. Master of science in aeronautical and mechanical engineering, Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, Brazil.

[20] Guimarães Neto, A. B. (2014). *Flight Dynamics of Flexible Aircraft Using General Body Axes: A Theoretical and Computational Study*. Doctor of science in aeronautical and mechanical engineering, Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, Brazil.

COPYRIGHT STATEMENT