DAMPER MODEL IDENTIFICATION USING HYBRID PHYSICAL AND MACHINE LEARNING BASED APPROACH

Michele Zilletti, michele.zilletti@leonardocompany.com, Leonardo Helicopters (Italy)

Ermanno Fosco, ermanno.fosco@leonardocompany.com, Leonardo Helicopters (Italy)

Abstract

In this paper the identification of a time domain model of a helicopter main rotor lead-lag damper is discussed. Previous studies have shown that lead-lag dampers have a significant contribution to the overall aircraft dynamics, therefore an accurate damper model is essential to predict complex phenomena such as, instabilities, limit cycles, etc. Due to the inherently nonlinear dynamics and the complex internal architecture of these components, the model identification can be a challenging task. In this paper, a hybrid physical/machine learning based approach, has been used to identify a damper model based on experimental test data. The model, called *grey box*, consists of a combination of a *white box*, i.e. a physical model described by differential equations and a *black box,* i.e. regression numerical model. The *white box* approximates the core physical behaviour of the damper while the *black box* improves the overall accuracy by capturing the complex dynamic not included in the *white box*. The paper shows that, at room temperature, the *grey box* is able to predict the damper force when either a multi-frequency harmonic or a random input displacement is imposed. The model is validated up to 20 Hz and for the entire damper dynamic stroke.

SYMBOLS

Symbol	Description	Units
β	Fluid compressibility	m^2/N
ζ	Average volume ratio between damper chambers	-
δ	Damper pressure difference between chambers	Pa
η	Damper fluid dynamic viscosity	kg/m/s
γ	Secondary discharge coefficient	m^-1
ω	Frequency	rad/s
ρ	Damper fluid density	kg/m^3
σ	Variance	-
θ_1	Variance of kernel function	-
$\boldsymbol{\xi}$	Optimisation vector	-
A	Damper displacement amplitude	m
A_p	Piston area	m
A_V	Relief valve area	m^2
c_1 and c_2	Relief valve damping	Ns/m
c_{exit}	Discharge coefficient	-
d	Damper orifice diameter	m
D_1 and D_2	Coefficient of laminar and exit loss	-
f	Damper force	N
$\Psi()$	Regression function	-
F_{c1} and F_{c2}	Relief valve preload forces	N
$\mathcal{GP}$	Gaussian Process	-
I	Identity matrix	-
J	Cost function	-
K	Damper stiffness	N/m
$\boldsymbol{k}()$	Covariance matrix function	-
k_1 and k_2	Relief valve stiffness	N/m
l_1	Length scale of kernel function	-
l_0	Damper orifice length	m
m	Number of GP conditioning values	-
$m()$	Mean function	-
m_V	Relief valve poppet mass	kg
NMSE	Normalised Mean Squared Errors	-
$p()$	Probability function	-
Q_0	Flow through the damper orifices	m^3/s
Q_v	Flow through the relief circuit	m^3/s
R	Relief valve coefficient loss	-
t	Time	s
T	Total number of training experiments	-
V	Chambers volume	m^3
w	Displacement of a particle of QPSO	-
x	Damper displacement	m
x_1 and x_2	Relief valve displacements	m
H_1, H, s_1 and s_2	QPSO hyper-parameters	-
x_{c1} and x_{c2}	Relief valve pre compression	m
x_V	Effective damper valve poppet displacement	m

1. INTRODUCTION

Lead-lag dampers are essential to guarantee the stability of helicopters as they increase the damping of the blade lead-lag motion [1]. Due to their significant contribution to the overall aircraft behaviour, an accurate model of their dynamics is essential to predict complex phenomena such as, instabilities, limit cycles, etc. [2], [3].

Lead-lag dampers have been widely studied and different architectures, such as hydraulic, elastomeric, to cite the most common ones, have been used in articulated rotors. The presence of the fluid and\or elastomeric materials have shown that the dynamics of these devices is inherently nonlinear

---

Presented at 47th European Rotorcraft Forum, Glasgow, UK, 7-9 September, 2021

and generally dependent on displacement, deformation rate, and temperature [4]–[9]. The classical way to model the damper dynamics is by linearization using the complex modulus. The damper is forced harmonically over a range of amplitudes, temperatures and frequencies (typically at the lead-lag resonance and\or at the blade passing frequency and higher order harmonics) and the complex modulus for each test condition is calculated from the imposed displacement, and measured force signals. The model could be as simple as a “look-up table” incorporating all the available test conditions, or more complex, involving a set of functions that could be curve-fit to all or part of the data set [10]. The approach used in this type of modelling is to calculate the damper’s linearized stiffness and damping for a given steady state flight condition to perform stability analyses. This approach does not allow to capture the inherent nonlinear behaviour of the damper and is inadequate to describe its transient dynamic response. Additionally, a linearized damper model, used in a helicopter dynamic simulation tool, may lead to inaccurate stability assessments of the aircraft and it would be unsuitable to predict nonlinear phenomena, such as limit cycles.

An alternative approach is to use a time domain model of the damper (often called physical model or *white box*), where differential equations (DE), describing the physical phenomena involved in the damper dynamics, are integrated over time. In this case, the model aims to predict the time history of the damper force for a given input displacement. References [6]–[8] show some examples of physical models which use simplified linear and nonlinear lumped parameter elements opportunely connected. The parameters are optimised by minimising the error between the model output and the experimental data. A more advanced model for a hydraulic damper is described by Eyres et al. in [11] where the piston-fluid interaction and the dynamics of relief valves are considered. In general, the accuracy of physical models depends on the complexity of the mathematical model, which, especially for some damper architectures involving both elastomeric materials and hydraulic components, may become prohibitively complicated and impractical to use. Also, these models are specific to the particular damper architecture [12].

The approach, considered in this paper, uses a white box model with a limited number of degrees of freedom and physical parameters to obtain a reasonable approximation of the core behaviour of the damper. The accuracy is then augmented with a numerical model (called *black box*), whose parameters are learnt from the experimental data using a machine learning general approximation approach. As the opposite of the *white box*, the fitted parameters will not have a physical meaning in this case. The combination of the *white* and *black box* is referred to as *grey box* model [13].

In this paper, firstly, the general dynamic behaviour of the damper based on the experimental data is described in Section 2. In Section 3 the derivation of the *grey box* model is described. The section is organised in two sub-sections dedicated to: the derivation of the differential equations of the *white box* model and their parameters optimisation; the description and fitting of the *black box* model using two machine learning regression approaches i.e.(Gaussian Process (GP) and Neural Network (NN)). The model prediction is finally compared against the experimental data and results are presented in Section 4. Conclusions are given in section 5.

2. LEAD-LAG DAMPER TESTS

The lead-lag damper under consideration uses the particular characteristics associated with the action of a high viscosity fluid and the deformation of an elastomeric material. The functionality of the damper can be represented by the simplified scheme in Figure 1. The movement of the piston forces the high-viscosity fluid through orifices due the pressure difference generated between the upper and lower chambers.

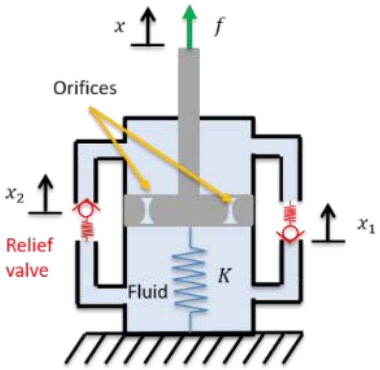

Figure 1: Simplified scheme of the damper.

A relief circuit, connecting the two chambers, allows the fluid passage if the pressure exceeds a threshold value. This is achieved by the use of relief valves calibrated by setting the preload on the valve springs. The reaction due to the deformation of the elastomeric material plus the compressibility of the high viscosity fluid is represented by the spring element K.

In order to gain a better understanding of the damper dynamics and collect the data needed for the model identification, characterisation tests have been carried out. The damper has been tested by grounding one of its terminal and imposing a displacement of the free terminal. Both the displacement and the force time histories have been recorded during the test. The simplest dynamic test consists in imposing a pure sinusoidal displacement signal x given by:

$$x = A\sin(\omega t) \tag{1}$$

where ω is the excitation frequency in rad/s and A is the displacement amplitude in mm.

As a preliminary step, the experimental data (i.e. the measured damper force f and the displacement x signals) have been pre-processed. Both signals have been filtered to opportunely remove the high frequency noise with a zero-phase digital filter in order to prevent phase distortions. The filter order has been set to the minimum value that guarantees a good noise reduction minimising the alteration of the high frequency damper dynamics. The velocity $\dot{x}$ has been then obtained by differentiating the displacement signal using a centred differences filter [14].

Figure 2 shows the displacement x, the velocity $\dot{x}$ and damper force f measured during a laboratory test when a pure sinusoidal displacement, given in equation (1), is imposed. As it can be noticed, the velocity signal is affected by noise. This is due to the differentiation step that amplifies the high frequency noise content. With the reference to A-B-C-D markers annotated in Figure 2, the dynamic behaviour of the damper can be described as follows:

A: The displacement and the force are at their maximum value and the velocity is almost zero;

A-B: As the displacement and the velocity decrease, the damper force generated by the piston-fluid interaction (the fluid is pushed through the orifices) becomes negative;

B: The relief valves are activated because the fluid pressure exceeds a threshold value;

B-C: Even though the relief valves are activated, (the fluid is free to flow in the relief circuit) the damper force continues to decrease, highlighting an elastic behaviour, until the displacement reaches a minimum;

C: The displacement and the force are at their minimum value and the velocity is zero thus the fluid pressure becomes smaller than the threshold value so the relief valves are deactivated;

C-D: As the velocity increases the force move from a negative to a positive value;

D: The relief valves open again as the fluid pressure exceeds the threshold value.

The same behaviour can be identified in the force vs displacement and force vs velocity diagrams shown in Figure 3. The velocity vs force diagram shows the presence of a nonlinear damping component and the finite area enclosed by the curve shows that the dynamic of the damper is characterized by a combination of stiffness and damping components. The presence of a stiffness component can be also noticed from the displacement vs force diagram.

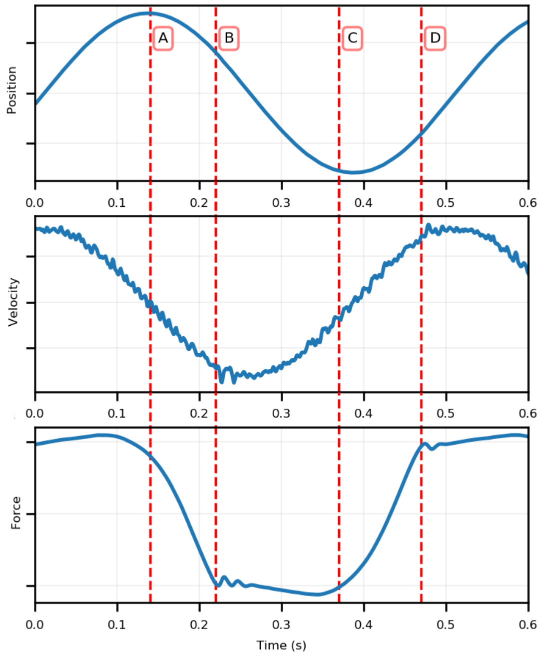

Figure 2: Measured displacement (top), velocity (mid) and force (bottom) acquired in an experimental test at single frequency.

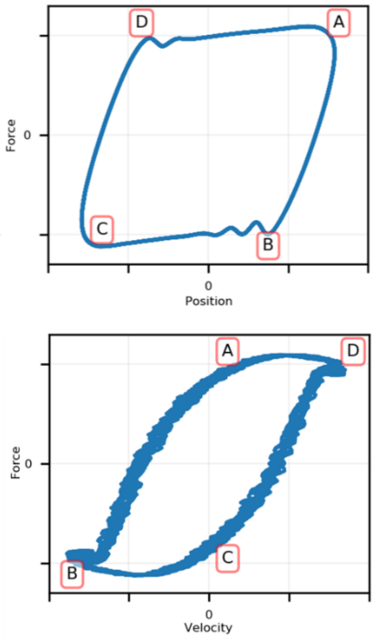

Figure 3: Force vs displacement (top) and force vs velocity (bottom) diagrams of the signal in Figure 2.

The full test campaign of the damper includes input displacement signals at different frequencies and amplitudes. A total of 96 experiments have been carried out. Figure 4 summarizes the amplitude and frequency content of the input displacement for each test. The bars in the top plot, shows the peak-to-peak amplitude while the blue markers represent the mean value; the bottom plot shows the frequency content. As it can be notice from the graph, some of the experiment labels on the x-axis are marked in bold to indicate that they are part of the training set of the *black box* model, as discussed in the following section. The full test campaign includes input displacement signals characterized by harmonics at the lead-lag frequency, ω_{lag}, the main rotor frequency 1/rev and two higher order harmonics, 2/rev and 3/rev at different amplitudes within the typical damper operating stroke.

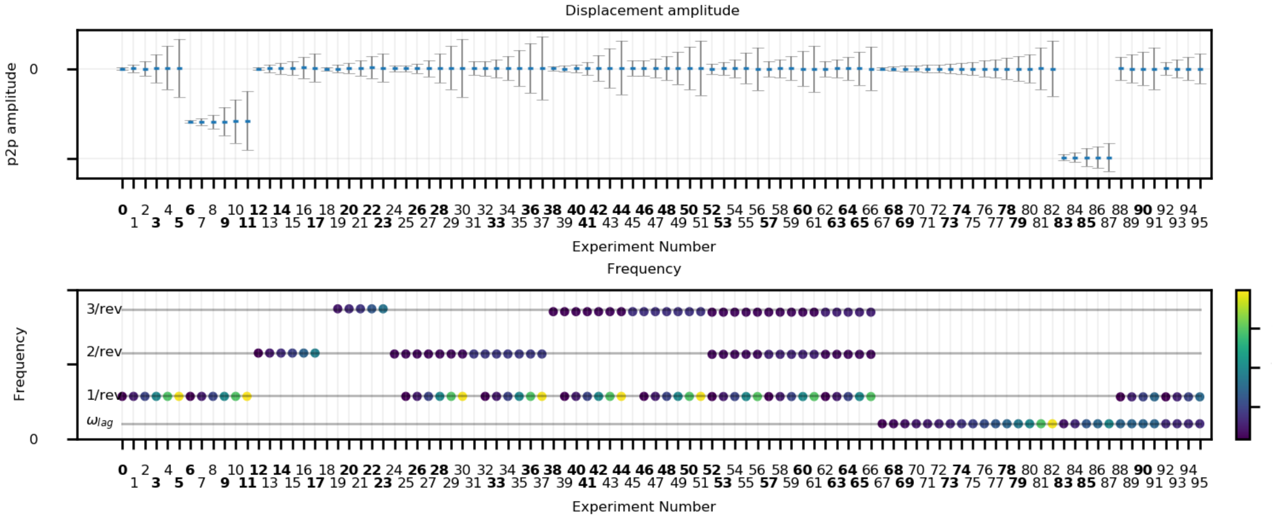

Figure 4: Peak-to-peak and mean amplitude (top plot) and frequency (bottom plot) of the input displacement for each experiment; colorbar show the amplitude of harmonic componets. The experiment numbers in bold are part of the training data set.

3. GREY BOX MODEL

Figure 5 shows the structure of the grey box model used for the estimation of the damper force starting from the input displacement $x(t)$ and velocity $\dot{x}(t)$. The displacement and velocity sample at instant t_i are passed to the *white box* to produce an estimate $\hat{f}_w(t_i)$ of the damper force integrating the differential equation describing the damper dynamics. The past n samples of the displacement $[x(t_{i-n}) \dots x(t_i)]^T$, velocity signal $[\dot{x}(t_{i-n}) \dots \dot{x}(t_i)]^T$ and the *white box* output $\left[\hat{f}_w(t_{i-n}) \dots \hat{f}_w(t_i)\right]^T$ are passed to the *black box* which produces the final estimate of the damper force $\hat{f}(t_i)$ accounting for the damper dynamic not capture by the *white box*. In the next subsections the derivation of the *white* and *black boxes* is described.

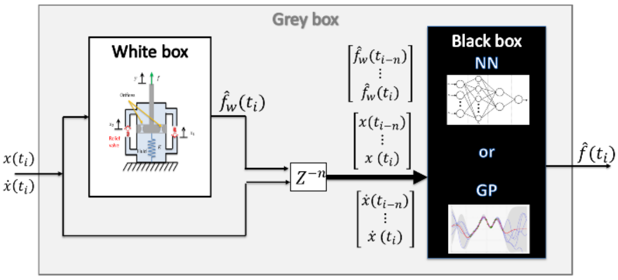

Figure 5: Grey box model scheme.

3.1. White box model

In this section, the differential equations (DEs) are derived and their parameters are optimised using the Quantum Particle Swarm Optimisation (QPSO) algorithm [15].

With reference to the scheme in Figure 1, the pressure difference between the upper and lower chambers δ, can be described by the following differential equation [11]:

$$\dot{\delta} = \frac{1+\zeta}{\beta \bar{V} \zeta}\left[A_p \dot{x} - \mathrm{sign}(\delta)(Q_0 + Q_V)\right], \tag{2}$$

where β is the fluid compressibility, $\bar{V}$ is the chambers mean volume and ζ is the average ratio between chambers. A similar approach has been used in the modelling of an automotive engine mounts as described in reference [16]. The total fluid volume flow induced by the piston movement $A_p\dot{x}$, where $\dot{x}$ is the piston velocity and A_p is the piston area, is equal to the fluid flow though the orifices, Q_0 plus the fluid flow through the relief circuit Q_V. The volumetric flow through the orifices Q_0 is a function of the pressure difference δ, and assuming a combination of laminar and turbulent flow characteristics, can be expressed as [11]:

$$Q_0 = \frac{-D_1 + \sqrt{D_1^2 + 4D_2|\delta|}}{2D_2}, \tag{3}$$

where D_1 and D_2 are the coefficient of laminar and exit loss respectively given by:

$$D_1 = \frac{128\eta l_0}{\pi d^4}, \qquad D_2 = \frac{c_{exit}\rho}{2\pi^2\left(\frac{d}{2}\right)^4}, \tag{4}$$

where η is the dynamic viscosity, l_0 and d are the orifice length and diameter respectively, c_{exit} is the discharge coefficient and ρ is the fluid density.

The volumetric flow rate through the relief valves Q_V is a function of δ and the valve effective poppet displacement x_V:

$$Q_V = R\frac{\gamma}{1+\gamma x_V}\, x_V\sqrt{|\delta|}\,, \tag{5}$$

where R and γ are the relief valve coefficient loss and the secondary discharge coefficient respectively. The valves poppet displacements x_1 and x_2, of the two relief valves, one for each flow directions, are modelled as a second-order dynamic system as:

$$m_V\ddot{x}_1 + c_1\dot{x}_1 + k_1(x_1 - x_{c1}) = A_V\delta\,, \tag{6}$$

$$m_V\ddot{x}_2 + c_2\dot{x}_2 + k_2(x_2 + x_{c2}) = A_V\delta\,, \tag{7}$$

Where m_v is the valve moving mass and A_V is the valve area. Also x_{c1} and x_{c2} are the valve pre compression displacements given by $x_{c1} = F_{c1}/k_1$ and $x_{c1} = F_{c2}/k_2$ where $k_1$and k_2 are the valves spring stiffnesses and F_{c1} and F_{c2} are the preload forces. Finally c_1 and c_2 are the damping coefficients. The effective poppet displacement x_v in equation (5) is given by:

$$x_v = \begin{cases} |x_1| \quad \text{if } \delta < -\dfrac{F_{c1}}{A_v} \text{ (valve 1 open)} \\ |x_2| \;\; \text{if } \delta > \dfrac{F_{c2}}{A_v} \text{ (valve 2 open)} \\ \quad 0 \quad \text{Otherwise} \end{cases} \tag{8}$$

The mechanism governing the elastomeric material deformation and fluid compressibility can be quite complex and dependent on environmental conditions such as temperature. In the current model, the elastic component has been approximated by a linear spring element in parallel with the piston. Even though more complicated differential equations may be required, the approach used here is to keep the *white box* as simple as possible and model the more complicated dynamics using a *black box*, as it will be discussed in the next section. Under this assumption, the overall white box force estimation $\hat{f}_w$ is given by the force generated by the fluid pressure acting on the piston area plus the stiffness component leading to:

$$\hat{f}_w = A_p\delta(\dot{x}) + Kx \tag{9}$$

where K is the linear stiffness coefficient of the spring element in parallel with the piston. The differential equation (9) describes the *white box* model able to provide a prediction of the damper force for an imposed damper displacement and velocity.

3.1.1 White box parameter optimisation

The majority of the physical parameters used to derive the white box can be obtained by design (e.g. mass, dimensions, etc.). However, for some of the parameters (e.g. stiffness, preload of the valves, effective piston area) it is difficult to have an accurate estimate because a direct measurement may be difficult to carry out or the nominal value may be affected by environmental factors such as temperature. In order to overcome this issue, some of the parameters have been tuned using an optimisation algorithm that minimises the discrepancy between the output of the white box and the measured damper force during the experiments.

The cost function $J(\xi)$ to be minimised is defined as the sum of the T Normalised Mean Squared Errors (NMSE) calculated for each τ experiment:

$$J(\boldsymbol{\xi}) = \sum_{\tau=1}^{\mathrm{T}} NMSE_\tau(\boldsymbol{\xi})\,, \tag{10}$$

where $NMSE_\tau$ is the Normalised Mean Squared Error (NMSE) of the τ-th experiment given by:

$$NMSE_\tau(\boldsymbol{\xi}) = \frac{100}{N\sigma_f^2}\sum_{i=0}^{N}\left(f_{e,\tau}(t_i) - \hat{f}_\tau^w(t_i,\boldsymbol{\xi})\right)^2\,, \tag{11}$$

where $f_{e,\tau}$ is the experimental recorded force, σ_f^2 is its variance and $\hat{f}_\tau^w$ is the output of the *white box* model. Also, t_i and N are the i-th time sample and the total number of samples respectively. Finally $\boldsymbol{\xi} = \left[F_{c1}, F_{c2}, K, A_p\right]^T$ is the vector of the selected tuning parameters. Since the time integration is computationally very demanding, it was unpractical, to include all the experiments time histories in the calculation of J in equation (10), therefore only some of the single frequency experiments with zero mean amplitude for a total number of $N = 8$ have been considered (specifically experiments 1, 4, 69, 68, 71, 74, 77 and 80). Finally, the optimal parameter vector $\boldsymbol{\xi}_{opt}$ can be expressed as:

$$\boldsymbol{\xi}_{opt} = arg\,\min_{\boldsymbol{\xi}} J(\boldsymbol{\xi}) \tag{12}$$

The algorithm used for the optimisation is a Quantum-behaved Particle Swarm Optimisation (QPSO) algorithm [15]. The QPSO method is based on the same principles as the regular particle swarm [17] but the dynamics of each particle is changed from the classic formulation to one where every particle is treated in a quantum manner [15]. The PSO algorithm is motivated by the flocking behaviour of flights of birds in search of food. Each particle (bird) k in the population (flock) is represented by a vector of its position and velocity ($w_k;\ \dot{w}_k$); the update rules for a generation are:

$$\begin{gathered}\dot{w}_k = \dot{w}_k + s_1 H_1 (w_k^* - w_k) + s_2 H_2 (w^* - w_k) \\ w_k = w_k + \dot{w}_k\end{gathered} \tag{13}$$

Where $\mathrm{w_k^*}$ is the is the best position experienced so far by the particle k, measured in terms of the cost function (10), and $\mathrm{w^*}$ is the best position of any particle. $\mathrm{s_1}$ and $\mathrm{s_2}$ are hyper-parameters and $\mathrm{H_1}$ and $\mathrm{H_2}$ are randomly generated. In the ‘quantum’ version of the PSO algorithm, the trajectory of the particles is dependent on, i) an attractor that is a random combination of the global best and previous best position for each particle, and ii) a term relating to the particles' potential fields.

Figure 6 shows the mean cost function value (red dashed line) across 10 particles and the best value (black solid line) during the convergence of the QPSO algorithm. The graph shows that all particles converge to an optimal cost function value after about 17 iterations.

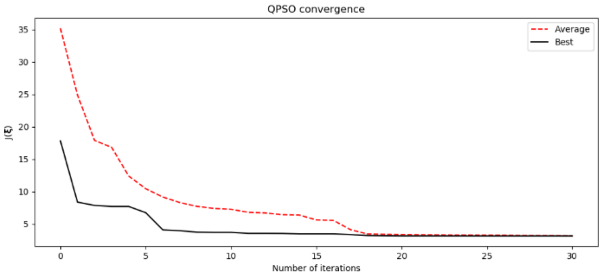

Figure 6: QPSO convergence. Mean value (red dashed line) across 10 particles and best value (black solid line) as a function of number of iterations.

After the parameters have been optimised, the NMSE has been calculated for all the 96 experiments and plotted in Figure 11 with the blue bars. The plot shows that NMSE is below 9 %. Figure 12 shows the measured time histories (blue line) and the output of the *white box* for experiments 1, 61 and 84 as an example. A comparison of the time histories highlights that the *white box* is able to capture the overall damper behaviour, however the prediction can be further improved by using a *black box*.

3.2. Black box model

The *black box* is implemented in the nonlinear autoregressive exogenous (NARX) framework, which can be described in the general regression form as:

$$\begin{gathered}\hat{f}_i = \Psi(\hat{f}_{i-1}, \dots, \hat{f}_{i-n_y}, \hat{f}_{i-1}^w, \dots, \hat{f}_{i-n_x}^w, \\ x_{i-1}, \dots, x_{i-n_x}, \dot{x}_{i-1}, \dots, \dot{x}_{i-n_x})\end{gathered} \tag{14}$$

i.e. regress the current system output on a range of past system inputs (displacement, x, velocity $\dot{x}$, output of the *white box* $\hat{f}^w$) and outputs $\hat{f}$. The function Ψ represents a numerical model, which in this paper has been derived using two different methods: Gaussian Process (GP) and Neural Network (NN). When fitting a NN regression model is good practice to divide the data in training, validation and test sets. In the first step, the training dataset are used to fit the model parameters. In a second stage, the hyper-parameters are tuned on the validation data set, which provides an unbiased evaluation of the model fit. In this case, the hyper-parameters are the number of past inputs and output samples (n_x in equation (14)) plus some other hyper-parameters specific to the particular method used as it will be discussed in the next two subsections. It has been chosen not to feedback the previous outputs y ($n_y = 0$ in equation (14)) to the *black box* as this may cause the accumulation of the prediction error at each step leading to a divergent output. Finally, the accuracy of the model is evaluated on the test dataset. The train data set, in this case, is composed by about 40% of randomly selected experiments from the 96 available (marked in bold on the x-axis in Figure 11). Of the selected experiments, only 50% of the samples have been included in the training set. The number of training samples is 1.47×10^4. The validation set is made of 20% of randomly selected samples from the remaining experiments. Test data set includes all other samples not part of the previous two groups. In the case of GP the Bayesian framework is robust against overfitting therefore only the hyper-parameters have to be learned from the validation set.

3.2.1 Gaussian process

The basic concept of a Gaussian process (GP) is to perform inference over functions directly, instead to inference over parameters of a function as it is done in a neural network regression, for example. A GP is a distribution over functions, which is conditioned on training data so that the most probable functions are the best fit to the data [18].

Let $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2 \dots \boldsymbol{x}_N]^T$ denotes a matrix of multivariate training inputs where the i-th row $\boldsymbol{x}_i = [\hat{f}_{i-1}^w, \dots, \hat{f}_{i-n_x}^w, x_{i-1}, \dots, x_{i-n_x}, \dot{x}_{i-1}, \dots, \dot{x}_{i-n_x}]$ in this case, and $\boldsymbol{f}$ denote the corresponding output vector of training measured damper forces. The input vector for a testing point will be denoted by the column vector $\boldsymbol{x}^*$ and the corresponding unknown output, which we aim to predict, is denoted by f^*. It is possible to place a Gaussian process prior over the nonlinear function in other words; the parametric function that fits the data is drawn from the Gaussian process defined as follow:

$$f(\boldsymbol{x}) = \mathcal{GP}(m(\boldsymbol{x}), \boldsymbol{k}(\boldsymbol{x}, \boldsymbol{x}')) \tag{15}$$

Where $m(\boldsymbol{x})$ is the mean function and $\boldsymbol{k}(\boldsymbol{x}, \boldsymbol{x}')$ is the positive define covariance matrix.

This GP will now generate many smooth/wiggly functions, and if the function to be predicted falls into this family of functions that GP generates, this is now a sensible way to perform nonlinear regression. The shape of the generated function can be set by choosing the covariance function used to generate the covariance matrix $\boldsymbol{k}$, which in this case is chosen to be:

$$\boldsymbol{k}(\boldsymbol{x}_m, \boldsymbol{x}_n) = \theta^2 \exp\left(\frac{1}{2l_1^2}(\boldsymbol{x}_n - \boldsymbol{x}_m)^2\right) \tag{16}$$

where θ and l_1 are two hyper parameters to be learnt from the training data.

One of the defining properties of the GP is that the joint probability of two finite set of data, for example the training set and the test set, is Gaussian multivariate. This can be used to condition the test set (unobserved) to the training set (observed) effectively fitting the GP model. Following the Bayesian approach and assuming the prior mean to be zero and a Gaussian noise model on the observation of variance σ^2, the joint probability of training and test data can be written as:

$$\begin{aligned} &\boldsymbol{k}(f, f^*) \\ &= \mathcal{N}\left(0, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) + \sigma^2\boldsymbol{I} & \boldsymbol{K}(\boldsymbol{X},\boldsymbol{x}^*) \\ \boldsymbol{K}(\boldsymbol{x}^*,\boldsymbol{X}) & k(\boldsymbol{x}^*,\boldsymbol{x}^*) + \sigma^2 \end{bmatrix}\right) \end{aligned} \tag{17}$$

Where $\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X})$ is the covariance matrix whose (i,j) element is given by $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$, I is the identity matrix, $\boldsymbol{K}(\boldsymbol{X},\boldsymbol{x}^*)$ is a column vector whose i element is given by $\boldsymbol{K}(\boldsymbol{x}_i,\boldsymbol{x}^*)$ and $\boldsymbol{K}(\boldsymbol{x}^*,\boldsymbol{X})$ is its transpose.

In order to use the above results, it is necessary to derive the probability of the test data conditioned on the train data which can be written as:

$$p(f^*|\boldsymbol{f}) = \frac{p(\boldsymbol{f}, f^*)}{p(\boldsymbol{f})} = \mathcal{N}(m(\boldsymbol{x}^*), k(\boldsymbol{x}^*,\boldsymbol{x}^*)) \tag{18}$$

Where $m(\boldsymbol{x}^*)$ and $k(\boldsymbol{x}^*,\boldsymbol{x}^*)$ are the posterior mean and posterior variance given by:

$$m(\boldsymbol{x}^*) = \boldsymbol{K}(\boldsymbol{x}^*,\boldsymbol{X})(\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{f} \tag{19}$$

And

$$\begin{aligned} k(\boldsymbol{x}^*,\boldsymbol{x}^*) = k(\boldsymbol{x}^*,\boldsymbol{x}^*) + \sigma^2 \\ - \boldsymbol{K}(\boldsymbol{x}^*,\boldsymbol{X})(\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) \\ + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{K}(\boldsymbol{X},\boldsymbol{x}^*) \end{aligned} \tag{20}$$

Equation (19) can be used to calculate the most probable prediction of the unknown damper force f^* given the training set while the confidence of the prediction is expressed by the variance in equation (20). One of the main limitation of Gaussian processes is their scalability to large datasets due to the inversion of the covariance matrix in equation (19). Stochastic Variational Gaussian Process (SVGP) has been designed to overcome this limit [19] and scale the GP method to larger data sets. The main idea behind SVGP is to approximate the true GP posterior with a GP conditioned on a smaller set of inducing test data set. This smaller set of M samples is representative of the full dataset. The benefit of using SVGP is that the complexity of the model reduces from $\mathcal{O}(N^3)$ where N is mumber of training samples for the classical GP to $\mathcal{O}(M^3)$. The accuracy of the method improves as M gets larger at the expenses of the computational effort.

The SVGP fitting was carried out using GPflow python package by maximising the evidence lower bound (ELBO) using Adam Optimiser [20]. The number of inducing values was set to $M = 200$ to compromise between computational effort and NMSE achieved on the validation data set.

The size of the mini-batch has been set to 1000 samples by compromising between the discrepancy between the ELBO calculated on the full data set and the average ELBO calculated across all the mini-batches and the computational time\memory. The top plot in Figure 7 shows the computational time taken to calculate the ELBO estimate for 20 mini-batches versus the mini-batch size. The graph shows that as the mini batch size increases the computation time increases. The bottom plot of Figure 7 shows the ELBO estimate versus the mini-batch size calculated 20 times. The plot shows that as the dimension of the mini-batch increases the ELBO estimate converge to the ELBO value calculated across the entire data set (ground truth).

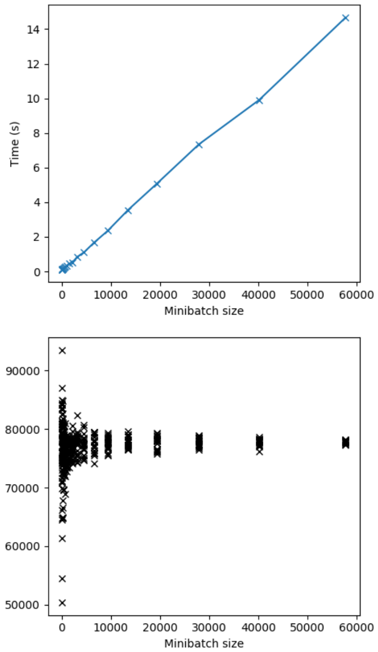

Figure 7: (top) computational time in seconds taken to estimate the ELBO of all mini-batch and (bottom) the ELBO estimate for each mini-batch vs the mini-batch size.

The size of the mini-batch has been set to 1000 as for mini-batch size greater than 1000 the spread of the ELBO estimate sharply decreases as shown in Figure 7 and is a good compromise for the computational time.

Figure 8 shows the histogram of the ELBO estimate calculated across 400 mini-batches of size 1000. Also shown is the ground truth (black line) and the mean ELBO estimate indicating that a good ELBO estimate can be achieved with a mini-batch dimension of 1000. Finally Figure 9 shows the estimate ELBO during the convergence of the SVGP optimisation with inducing value of 200 and mini-batch size 1000.

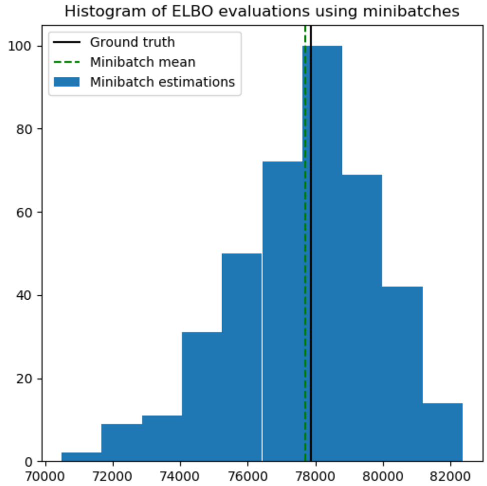

Figure 8: Histogram of ELBO estimate calculated across 400 mini-batches of size 1000 samples.

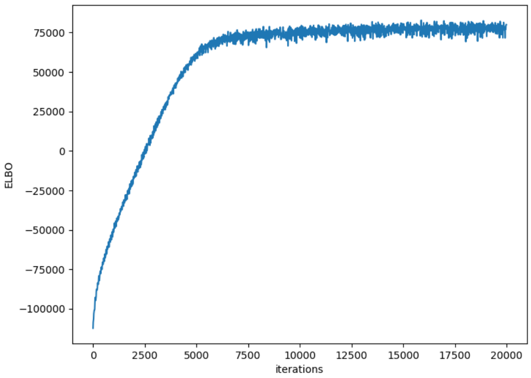

Figure 9: Convergence of SVGP.

3.2.2 Neural network

Neural networks (NN) are commonly used regression models, which use parametric forms of base nonlinear functions whose parameter values are optimised during training. There are different types of neural network, ranging from feed-forward NN to more complex types such as recurrent, convolutional, and others. The choice of the NN type depends on the application under consideration. In this study because the *white box* provides a well correlated signal to the measured damper force, it has been observed that a relatively simple architecture, such as feedforward NN, is able to give satisfactory accuracy. The NN architecture (number of neurons and hidden layers) was optimised using the validation data set. A network with a small number of neurons and hidden layers is, in general, not able to model complex nonlinear behaviours and it would result in a high NMSE on the training data set (high bias). On the opposite, a very high number of neurons and hidden layers would over fit the training data but it will not generalise on the test data set (high variance) [18]. The NN used for this application has one input layer of 128 neurons, three fully connected hidden layers of 64, 32 and 16 neurons respectively and a single output neuron. The activation function used is the ReLU function. The results relative to the architecture optimisation are not reported here for brevity. The NN training has been performed using Python Keras [21].

3.2.3 Black box fitting

In order to find the optimal number of input lags, the model has been trained on the training data set and the NMSE has been calculated on the validation set. Figure 10 shows the NMSE for the training and validation set as a function of the inputs lag number for the GP (left plot) and NN (right plot) regressions. In both cases, an increase of the number of lags, reduces the NMSE.

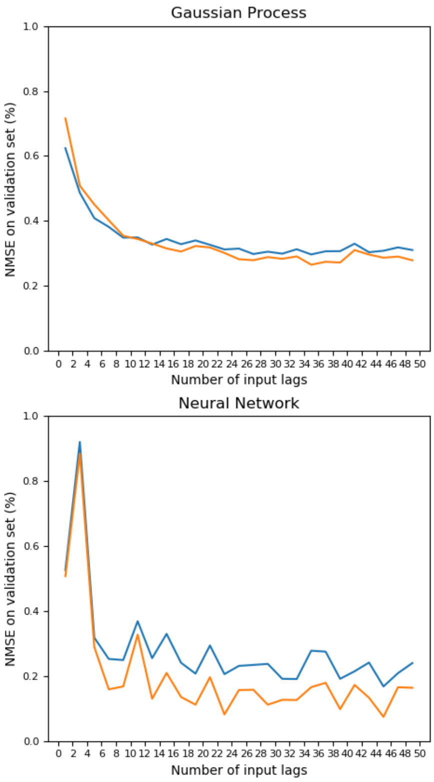

Figure 10: NMSE calculated on the train and validation set as a function of the number of input lags for: GP (left plot) and NN (right plot).

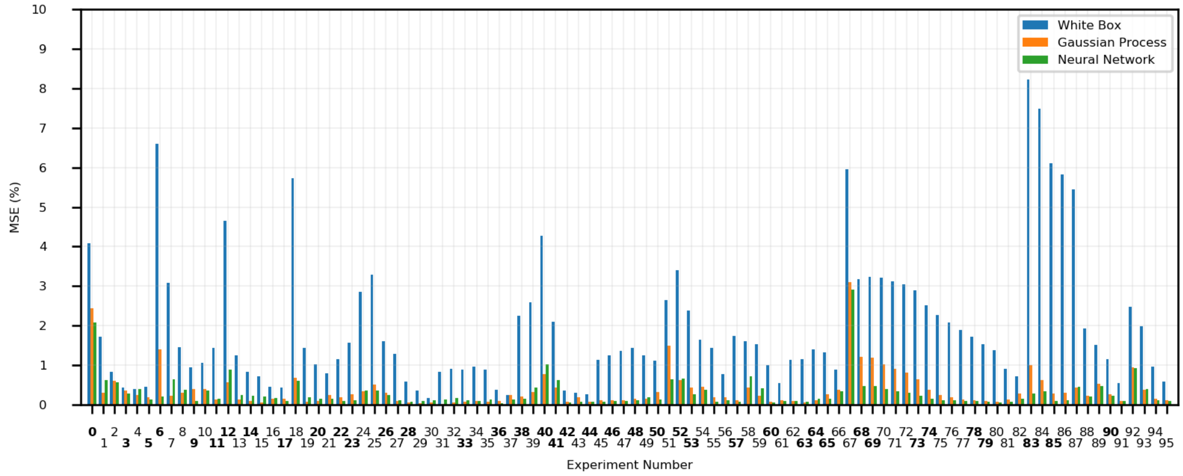

Figure 11: NMSE (%) for each experiment obtained with the white box (blue), grey box using GP (orange) and NN (green). Training set experiments marked in bold.

The small discrepancy of the NMSE calculated on the training and validation data sets indicates that the regression model is able to generalise the prediction on unseen data and therefore is free of overfitting. The number of lags has been fixed to 15 since, as shown in Figure 10, after an initial steep reduction the NMSE curve flattens out. Keeping the number of lags as low as possible, reduce the training computational effort.

4. RESULTS

After the hyper-parameters have been optimised, and the model has been trained, the NMSE has been calculated for each of the 96 experiments.

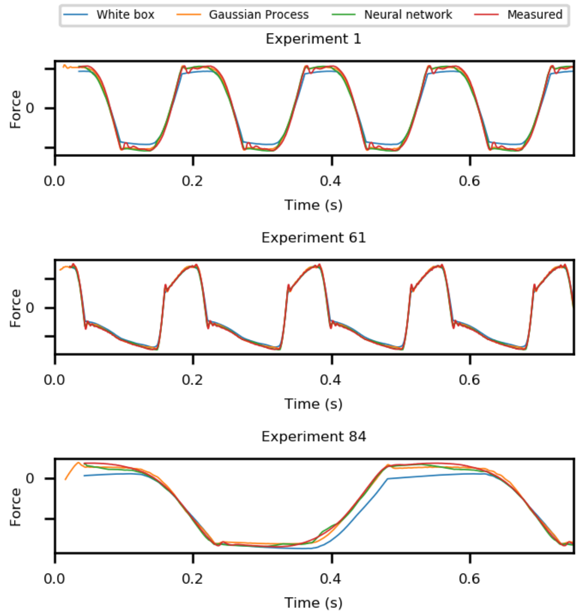

Figure 12: Comparisons damper force time histories of experiment 1 (top plot), 61 (centre plot) and 84 (bottom plot). White box (blue line), grey box using GP (orange line), grey box NN (green line) and measured (red line).

Figure 11 shows that the NMSE improves the *white box* prediction bringing the NMSE below 3%. In particular, the NMSE values obtained for the GP (orange bars) and NN (green bars) are comparable. The NN seems to slightly outperform the GP regression and has the advantage to be computationally less demanding.

To better appreciate the improvement obtained with the *black box*, the damper force prediction obtained with the GP (orange line) and NN (green line) regressions have been plotted in Figure 12 for three selected experiments. As discussed in Section 3, also the *white box* prediction (blue line) and the measured force (red line) are plotted for comparison. The plot shows that the use of the *black box* improves the accuracy by predicting the dynamics not captured by the *white box* model.

4.1. Random input displacement test

In order to further validate the model, the damper has been tested imposing a random displacement input of 30 seconds with a cut-off frequency of 20 Hz.

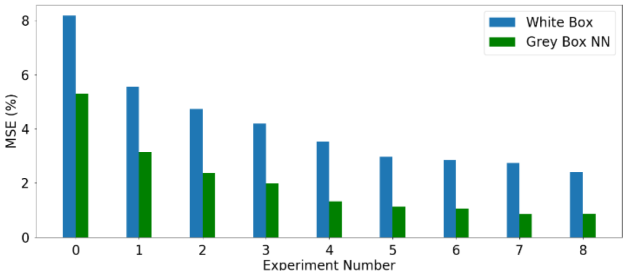

Figure 13: NMSE (%) for each experiment with random input obtained with the white box (blue), grey using NN (green). Model trained on multi-harmonic displacement of section 2.

Nine tests have been carried out with increasing root mean squared (rms) values of the input signal to cover the full stroke of the damper. Figure 13 shows the NMSE obtained either with *white box* (blue bars) or *grey box* (greed bars) models. The plot shows that the NMSE of the *grey box* is below 5%. This result shows that the model gives a good estimate of the damper force also for random input displacements not used in the model fitting (white box parameter optimisation process and black box training) which was done based on harmonic displacements only.

The accuracy of the model improves as the input displacement rms gets larger. This is due to the fact the majority of the tests used for the model fitting were carried out at high displacement amplitude which caused the opening of the relief valves. It should be also mentioned that the test rig showed some limitations when imposing low amplitude displacements due to the limits imposed by the hydraulic actuator of the test rig and the presence of friction phenomena.

Figure 14 shows 3 seconds of the time history measured force (red line), the *white box* (blue line) and the *grey box* prediction (green line) for comparison.

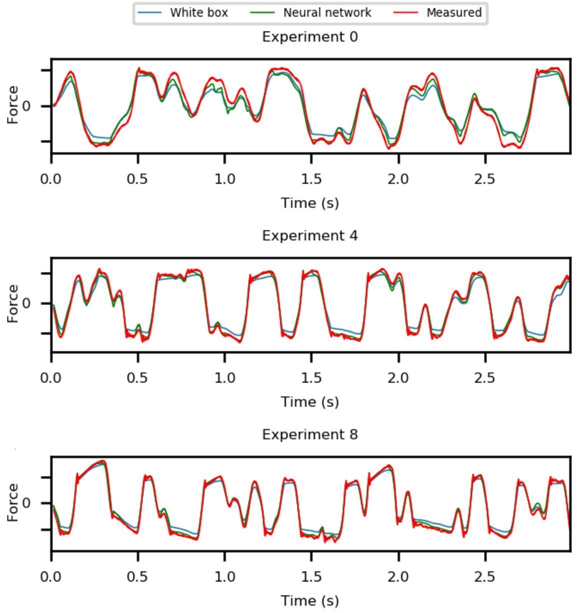

Figure 14: Comparisons damper force time histories of random input experiment 0 (top plot), 4 (centre plot) and 8 (bottom plot). White box (blue line), grey box using NN (green line), and measured (red line).

5. CONCLUSIONS

This paper has shown the identification of a time domain dynamic model of a helicopter main rotor lead-lag damper. The lead-lag damper under consideration uses the peculiar characteristics associated to the action of a high viscosity fluid and the deformation of an elastomeric material, which makes the damper dynamics inherently nonlinear.

The modelling approach used is the so-called *grey box* model, i.e. a combination of a physical model (*white box*) and a numerical regression model (*black box*). The *white box*, consisting of a set of differential equations, describes the core dynamic of the damper. A QPSO algorithm has been used to optimise the parameters of the *white box* based on experimental test data obtaining a NMSE below 9%. A *black box*, consisting of a general machine learning approximation approach, was then used to improve the overall model accuracy capturing the complex dynamics not described in the *white box*. Two regression methods for the fitting of the *black box* have been compared: Gaussian Process and Neural Network. The two methods give similar NMSE however, the NN is the preferred option for its lower computational effort.

The *grey box* damper force prediction obtained gave a NMSE below 5% at room temperature both for multi-harmonic and random input displacements for the entire damper stroke.

6. REFERENCES

[1] W. Johnson, *Rotorcraft Aeromechanics*. Cambridge University Press, 2013.

[2] G. Quaranta, V. Muscarello, and P. Masarati, “Lead-Lag Damper Robustness Analysis for Helicopter Ground Resonance,” *J. Guid. Control Dyn.*, vol. 36, no. 4, pp. 1150–1161, Jun. 2013.

[3] B. Titurus and N. Lieven, “Integration of Hydraulic Lag-Damper Models with Helicopter Rotor Simulations,” *J. Guid. Control Dyn.*, vol. 33, no. 1, pp. 200–211, Jan. 2010, doi: 10.2514/1.41961.

[4] M. I. Wallace, D. J. Wagg, S. A. Neild, P. Bunniss, N. A. J. Lieven, and A. J. Crewe, “Testing coupled rotor blade–lag damper vibration using real-time dynamic substructuring,” *J. Sound Vib.*, vol. 307, no. 3, pp. 737–754, Nov. 2007.

[5] D. P. McGuire and Lord Corporation, “Fluidlastic Dampers and Isolators for Vibration Control in Helicopters,” presented at the AHS International Forum 50, May 1994.

[6] B. Panda, E. Mychalowycz, and F. J. Tarzanin, “Application of passive dampers to modern helicopters,” *Smart Mater. Struct.*, vol. 5, no. 5, pp. 509–516, Oct. 1996.

[7] D. L. Kunz, “ELASTOMER MODELLING FOR USE IN PREDICTING HELICOPTER LAG DAMPER BEHAVIOR,” *J. Sound Vib.*, vol. 226, no. 3, pp. 585–594, Sep. 1999.

[8] P. J. Jones, D. D. Russell, and D. P. McGuire, “Latest Developments in Fluidlastic (R) Lead-Lag Dampers for Vibration Control in Helicopters,” presented at the AHS 59th Annual Forum Proceedings, Phoenix, Arizona, May 2003, vol. 1.

[9] E. C. Smith, G. A. Lesieutre, J. T. Szefi, C. Marr, and T. P. S. University, “Time Domain Fluidlastic Lag Damper Modeling,” presented at the AHS International Forum 62, May 2006.

[10] F. F. Felker, B. H. Lau, S. McLaughlin, and W. Johnson, “Nonlinear Behavior of an Elastomeric Lag Damper Undergoing Dual-Frequency Motion and its Effect on Rotor Dynamics,” *J. Am. Helicopter Soc.*, vol. 32, no. 4, pp. 45–53, Oct. 1987.

[11] R. D. Eyres, A. R. Champneys, and N. A. J. Lieven, “Modelling and Dynamic Response of a Damper with Relief Valve,” *Nonlinear Dyn.*, vol. 40, no. 2, pp. 119–147, Apr. 2005.

[12] K. Worden, D. Hickey, M. Haroon, and D. E. Adams, “Nonlinear system identification of automotive dampers: A time and frequency-domain analysis,” *Mech. Syst. Signal Process.*, vol. 23, no. 1, pp. 104–126, Jan. 2009.

[13] K. Worden *et al.*, “On evolutionary system identification with applications to nonlinear benchmarks,” *Mech. Syst. Signal Process.*, vol. 112, pp. 194–232, Nov. 2018.

[14] K. Worden, *Nonlinearity in Structural Dynamics: Detection, Identification and Modelling*. CRC Press, 2019.

[15] Jun Sun, Bin Feng, and Wenbo Xu, “Particle swarm optimization with particles having quantum behavior,” in *Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No.04TH8753)*, Jun. 2004, vol. 1, pp. 325-331 Vol.1.

[16] A. Kyprianou, J. Giacomin, K. Worden, M. Heidrich, and J. Bocking, “Differential evolution based identification of automotive hydraulic engine mount model parameters,” *Proc. Inst. Mech. Eng. Part J. Automob. Eng.*, vol. 214, no. 3, pp. 249–264, Mar. 2000, doi: 10.1243/0954407001527402.

[17] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in *MHS’95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Oct. 1995, pp. 39–43.

[18] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.

[19] J. Hensman, N. Fusi, and N. D. Lawrence, “Gaussian Processes for Big Data,” 2013.

[20] A. G. de G. Matthews *et al.*, “GPflow: A Gaussian process library using TensorFlow,” *J. Mach. Learn. Res.*, vol. 18, no. 40, pp. 1–6, Apr. 2017.

[21] F. Chollet and et al., *Keras, https://keras.io*. 2015.