



**HAL**  
open science

# Surrogate modelling for an aircraft dynamic landing loads simulation using an LSTM AutoEncoder-based dimensionality reduction approach

Michele Lazzara, Max Chevalier, Michele Colombo, Jasone Garay Garcia, Corentin J. Lapeyre, Olivier Teste

## ► To cite this version:

Michele Lazzara, Max Chevalier, Michele Colombo, Jasone Garay Garcia, Corentin J. Lapeyre, et al.. Surrogate modelling for an aircraft dynamic landing loads simulation using an LSTM AutoEncoder-based dimensionality reduction approach. *Aerospace Science and Technology*, 2022, 126, pp.107629. 10.1016/j.ast.2022.107629 . hal-03840028

HAL Id: hal-03840028

<https://ut3-toulouseinp.hal.science/hal-03840028>

Submitted on 22 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Surrogate modelling for an aircraft dynamic landing loads simulation using an LSTM AutoEncoder-based dimensionality reduction approach

Michele Lazzara<sup>a,b,\*</sup>, Max Chevalier<sup>b</sup>, Michele Colombo<sup>a</sup>, Jasone Garay Garcia<sup>d</sup>, Corentin Lapeyre<sup>c</sup>, Olivier Teste<sup>b</sup>

<sup>a</sup>*Airbus Operations, 316 Route de Bayonne, 31060 Toulouse, France*

<sup>b</sup>*IRIT, Cr Rose Dieng-Kuntz, 31400 Toulouse, France*

<sup>c</sup>*CERFACS, 42 avenue Gaspard Coriolis, 31057 Toulouse, France*

<sup>d</sup>*Airbus Operations GmbH, Kreetzslag 10, 21129 Hamburg, Germany*

---

## Abstract

Surrogate modelling can alleviate the computational burden of design activities as they rely on multiple evaluations of high-fidelity models. However, the learning task can be adversely affected by the high-dimensionality of the system, complex non-linearities and temporal dependencies, leading to an inaccurate surrogate model. In this paper we present an innovative dual-phase Long-Short Term Memory (LSTM) Autoencoder-based surrogate model applied in an industrial context for the prediction of aircraft dynamic landing response over time, conditioned by an exogenous set of design parameters. The LSTM-Autoencoder is adopted as a dimensionality-reduction tool that extracts the temporal features and the nonlinearities of the high-dimensional dynamical system response, and learns a low-dimensional representation of it. Then, a Fully Connected Neural Network is trained to learn the simplified relationship between the input parameters and the reduced representation of the output. For our application, the results demonstrate that our LSTM-AE based model outperforms both Principal Component Analysis and Convolutional-Autoencoder based surrogate models, in predicting the parameters-dependent high-dimensional temporal system response.

---

\*Corresponding author

*Email address:* [michele.lazzara@airbus.com](mailto:michele.lazzara@airbus.com) (Michele Lazzara)

*Keywords:* Surrogate modelling, Deep Learning, Dynamical systems, Dimensionality-reduction, Multivariate time-series, Dynamic Landing

---

## Nomenclature

### Abbreviations

A/C	Aircraft
DoE	Design of Experiments
DSE	Design Space Exploration
IQs	Interesting Quantities
SM	Surrogate Model

### Acronyms

AE	Auto-Encoder
CNN	Convolutional Neural Network
FCNN	Fully Connected Neural Network
LSTM	Long Short-Term Memory
PCA	Principal Component Analysis
POD	Proper Orthogonal Decomposition

### Symbols

$\mathbf{x}_i$	vector of system parameters for the $i$ -th DoE sample
$\mathbf{Y}_i$	multivariate time series and dynamical response for the $i$ -th DoE sample
$\mathbf{z}_i$	vector of the reduced manifold space
$N_T$	number of time steps
$N_x$	number of system parameters
$N_y$	number of IQs time series
$N_z$	dimension of the latent space

## 1. Introduction

The interweave between computer simulations, experimental data, and data-driven methods is now one of the building blocks of the modern engineering. Data-driven models expressing the behaviour of a complex system are the foundation of Surrogate Models (SMs) [1]. Surrogate modelling aims at providing a simple, fast and robust engineering instrument able to emulate the input-output relation of high-fidelity models, and thus it offers an efficient alternative to costly processes as computational simulations or test campaigns. Indeed, multi-departmental physical models lack the flexibility activities in which the testing of different parallel scenarios is desired. Consequently, SMs can accelerate early-design engineering tasks such as design optimization [2, 3], design space exploration [4, 5], uncertainty analysis [6] and sensitivity analysis [7, 8]. Design Space Exploration (DSE) based on SMs can play a relevant role in the design phase of the aircraft development life-cycle. In the airframe loads sizing context, DSE based on SMs allows quickly addressing changes in loads envelope with respect to considered changes of the design space. The methodology presented in this paper has been developed in the context of a need to accelerate the loads calculations in a set of given business scenarios. In particular: *(i)* early design phase to understand the impact of the different design parameters on the behaviour of the loads distributions; *(ii)* understand the potential loads exceedances due to design changes between the different maturity gates of the A/C design phase. There is no application foreseen outside of this context, at least for the time being. It could be possible to think of other use cases, as detection of flight mechanics simulation anomalies or airfoil aerodynamics prediction [9, 10] in high-dimensionality, depending on the type of initial parameters that are changed, whether they are pure design parameters or of any other nature.

The data-driven SM is not supposed to completely replace the physical models for the aircraft sizing, but rather it represents a tool to support decision-making in early-design phase. In this framework, using data-driven SMs guarantees adequate speed in terms of lead time, early optimum design and uncertainty

30 management, which are top priorities in aerospace industry [11, 12].

Despite the increasingly widespread use of data-driven SMs, a key problem rarely covered relates to the cases where inputs and outputs of a system are high dimensional objects [13, 14, 15], and of different nature [16, 17]: building a surrogate model of a highly parametrized nonlinear dynamical system, as the  
35 ground loads simulation model of a flexible aircraft, still remains an open question.

In this paper, SMs are designed to estimate the deterministic function  $f$  which maps the system parameters  $\mathbf{x}$  to the parameter-dependent response  $\mathbf{Y}$ , i.e.  $f : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y \times N_T}$ , where  $N_x$  represents the number of input parameters,  $N_y$   
40 the number of output variables, and  $N_T$  the number of time steps. The time-series  $\mathbf{Y}$  reflects the dynamic response of the system, whose behaviour over time depends on an initial state defined by a set of the design parameters  $\mathbf{x}$  (Fig.1). Thus, the SM should be able to predict a multivariate time-series  $\mathbf{Y} \in \mathbb{R}^{N_y \times N_T}$ , from a vector of constant scalar values  $\mathbf{x} \in \mathbb{R}^{N_x}$ . Using data-driven techniques  
45 to estimate the temporal evolution of a dynamical system, either from different initial conditions [18] or system parameters [19] has become an important topic in the last years, as demonstrated by several studies in fluid mechanics applications [8, 18, 20, 21, 22, 23, 24, 25].

Classical data-driven techniques [26] can be severely affected by high-dimensionality  
50 in two ways: directly, with a huge number of training samples, and indirectly by an increasing number of predictions that must be performed, leading to the well-known *curse of dimensionality* [27] issue.

Dimensionality reduction techniques are often applied to alleviate the problem of the high-dimensionality of the system by mapping the original problem space  
55 to a suitable lower dimensional space. This transformation is employed to construct the SM directly in the reduced space, where the input-output mapping can be easier to learn, thereby providing a more accurate surrogate model.

The combination of linear reduction methods such as Principal Component Analysis (PCA) [28], also known as Proper Orthogonal Decomposition (POD)  
60 [29], with data-driven machine learning models has resulted in relevant ap-

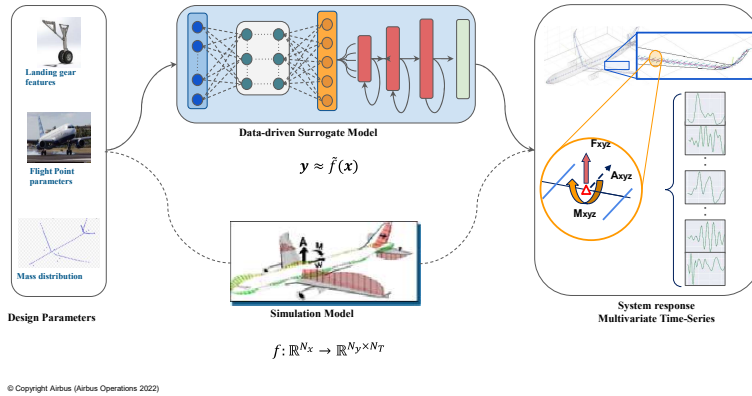


Figure 1: Surrogate modelling for computational efficient estimation of aircraft loads for dynamic landing.

proaches for surrogate modelling of large-scale dynamical systems. This approach shows its limitation when dealing with nonlinear problems, since PCA relies on learning linear projections as transformation of the data. To alleviate the linearity constraints, unsupervised learning techniques, such as Kernel-PCA [30, 31], IsoMap [32] or diffusion maps algorithm [33] are considered as nonlinear dimensionality reduction methods. However, besides the "pre-image" problem, they are often not capable of outperforming PCA [34, 35].

A more recent solution is represented by Auto-Encoders (AEs) [36]. An AE is a multi-layered neural network, whose "bottleneck" configuration forces to identify the essential attributes of the input data, leading to an embedded representation in a lower-dimensional space called "latent space". The attractive properties of AEs are: (i) the reduction mapping is learned jointly with the inverse mapping; (ii) as data dimensionality is reduced by multiple nonlinear transformation, AEs are able to describe the underlying nonlinear structure of the data. Convolution AEs [37, 20] have been proven to be efficient in surrogate modelling applications dealing with high dimensional and spatially distributed data, e.g. spatially discretized partial differential-equation models, typical of

fluid mechanics applications. Even if Convolutional Neural Networks (CNNs) are efficient for predicting time series [38, 39], they are not able to get both

80 short and long temporal correlations in situation where spatial neighbouring information is poor, because they treat time series as static objects. "Long short-term memory" (LSTM) [40] layers for the AE architecture have been selected for this study after a careful investigation. LSTM network is a system-identification based method, capable of capturing short-long time-delayed effects

85 [9] and thus dependency across time sequences, which are automatically memorized and embedded in the model. The same effect could be accomplished by a deep feedforward multi-layer network but that would require collecting the input vector across time and increasing the network size to look at long sequences, making its use harder or even intractable. Moreover, the characteristics

90 of our 1D time-series data does not justify the implementation of ConvLSTM [41]. Indeed, *(i)* they do not present properties of fluid flow simulations, or 2D image-like format simulations, where high-order spatial modes can be easily captured; *(ii)* although ConvLSTM filter sharing operation can capture spatiotemporal patterns, determining the right kernel dimension and the stride

95 step of convolutional operation is not a straightforward task, even tougher in an autoencoder setting where the configuration of the layers must be adjusted by hand. Furthermore, using ConvLSTM would increase the number of parameters proportionally to the filter window size with respect to LSTM, and finally would require a pre-study to find the optimal size of sub-sequences to being processed

100 by ConvLSTM. Instead, short and long term temporal dependencies, which are widely present in the dynamic landing simulations, are automatically learnt by LSTM network with relative effort on tuning the hyper-parameters. The proposed methodology reflects the will of the authors to provide an enhanced surrogate model for dynamic loads calculation. In aircraft loads modeling, POD

105 or PCA assisted surrogate modeling [42, 43] and Singular Value Decomposition (SVD) with Gaussian Process Regressor [44, 45] have been widely used to create approximate models of time-varying load envelopes, showing interesting results for uncertainty analysis. However the use of linear, 'variance-based', dimension-

ality reduction techniques is deficient with respect to the capability to identify  
110 non-linear temporal dependencies of correlated loads. Furthermore, the strategy  
'one regressor-one coefficient' to build local models seems infeasible in a high-  
dimensional context. Instead, a unique surrogate model would better capture  
the global inter-correlations among multiple variables, essential for the predic-  
tion of 1D and 2D envelopes.

115 In this paper a two-step learning strategy is proposed to tackle the shortcomings  
of previous data-driven techniques for loads prediction and fill the applicability  
gap with methods implemented in fluid mechanics applications.

A Recurrent Neural Network AutoEncoder based on LSTM neurons is adopted  
to extract the temporal features of a dynamic system response, encoding them  
120 in a reduced manifold. A Fully Connected Neural Network (FCNN), known  
to be an universal approximator [46] and to scale well with a large number  
of training samples, is then trained to map the input parameters space to the  
lower-dimensional representation of the original space, i.e. its latent represen-  
tation. This dual-step approach intrinsically produces more sources of error  
125 (two mapping errors and error propagation), compared to a singular model  
strategy. Nevertheless, dimensionality reduction may provide a more accurate  
model that has a lower error than a single-step surrogate approach, considering  
the facilitated mapping relation between inputs and low-dimensional outputs  
representation [4]. The combination of LSTM neurons and AE bottleneck ar-  
130 chitecture allows capturing the inter and intra dimensional nonlinear temporal  
correlations of the multivariate time-series. This information is embedded in  
a lower-dimensional space and used for the construction of the SM, allowing  
for an easier training task and consequently providing better interpolation and  
generalization capabilities.

135 The contributions of this work are three-fold:

- A novel strategy to build a SM that combines LSTM-AE and FCNN to  
predict a high-dimensional temporal correlated multivariate time-series  
conditioned by an exogenous set of constant parameters.



- Make a bridge between the studies on data-driven surrogate models for fluid mechanics application and data-driven methodologies used for the estimation of correlated loads envelopes.
- To the best of our knowledge, this is the first work on adopting purely deep learning techniques to predict dynamic aircraft correlated loads in a large-scale industrial simulator context.

This paper is organized as follows. Section 2 reviews the literature related to surrogate modelling and their application combined with data-driven dimensionality reduction techniques (Section 2.2). The section continues with a theoretical background on LSTM autoencoders (Section 2.3). In Section 3, the problem setting is introduced and the proposed approach described. The experiments and the discussion of the results are presented in Section 4. Finally, Section 5 concludes the paper with a brief summary and future research extensions.

## 2. Background

Our work is built upon previous models based on dimensionality reduction techniques for surrogate modelling purposes. In this section, data-driven surrogate modelling techniques are briefly introduced. Then a review is given on the use of hybrid PCA-based neural network model, as well as surrogate modelling based on AutoEncoders application. Finally, a theoretical background on LSTM autoencoder is provided, supporting the choice of using such approach in this research.

### 2.1. Surrogate Modelling

Surrogate modelling techniques aim at approximating the function relating a set of independent input parameters to a set of parameters-dependent variables, by exploiting the information within a dataset. The so-build SM is supposed to learn and interpret the existing high-fidelity solutions for predicting the response of the system for a new, previously unobserved input, in an agile and fast way. Data-driven surrogate modelling [47] does not need to explicitly formulate

the physics mathematically, but rather it learns physical information only from data gathered from previous high-fidelity experiments.

Traditional data-driven surrogate methods are powerful and flexible tools to quickly investigate several points in the design space. Polynomial chaos expansion regression [48, 49], Gaussian process regression (GPR) [50, 51] and fully-connected neural networks (FCNNs) [52, 53] have been widely used in aerodynamics [54, 55] and aircraft loads estimation [56, 57, 58]. These methods are characterized by versatility, low evaluation cost and easy handling. However, their learning task suffers the high-dimensionality of the problems for both the size of the training samples and the number of variables to predict. For example, in the case of Gaussian Process, computing the inverse of the covariance becomes challenging when the dataset is too large. This limits the maximum number of training points that the method can handle, which in turn limits the accuracy that can be achieved when many training points are available [13, 59].

## 2.2. Dimensionality reduction and Surrogate Models

Classical data-driven models can suffer from the high-dimensionality of the learning problem, where the dimensionality refers to the number of the overall values of a single system observation to be processed by the model. Learning about interdependencies of system state between input and output space represents itself a difficult task. Considering the temporal dimension makes the learning problem even harder: the input-output mapping capability is impacted by the increasing number (by a factor  $N_T$ ) of the overall values (e.g., temporal observations of physical quantities), and by the temporal dependencies among variables. To solve this problem, dimensionality reduction techniques [34] have been applied to produce lower dimensional data containing sufficient information to enable an easier learning task, leading thus to a more accurate surrogate. In the next section, prior research on surrogate models build with a PCA (or POD)-based approach and its non-linear derivatives is reviewed. Then, the use of convolutional autoencoder-based dimensionality reduction for surrogate modelling is investigated. This review, in addition to providing the reader with

a background on the use of dimensionality reduction for surrogate modelling, highlights also the disadvantages in using the mentioned techniques for our case of study.

200 *2.2.1. Hybrid PCA-based model*

A widespread family of approaches relies on the projection of data in a reduced subspace spanned by a set of orthonormal basis vectors and the corresponding modal coefficients. Particularly, such approach consists in training the machine learning model to map the input parameters to the reduced coordinates  
205 associated to the truncated basis functions capturing a certain amount of variance (energy) from data. In [8], a data-driven surrogate modelling framework was developed to emulate spatio-temporal gaseous and spray fields using Gaussian Process techniques over the POD-reduced numerical simulation data. In [60, 61] the authors propose a non-intrusive reduced basis method, where a GPR  
210 is used to approximate the POD projection coefficients. For the prediction of steady turbulent aerodynamic fields, local reduced-order models have been built for each input parameter subspaces using POD coupled with GPR [7]. In [44], a singular regressor is used for each column of the reduced matrix containing the left singular vectors resulting from the application of Singular Value Decomposition (SVD) to the time histories of the observable outputs. Similarly, in [62]  
215 Kriging-based surrogate models for each specific sensor are performed over the Power Spectrum Density diagrams of the time-varying signals.

The mapping between the input parameters and the basis expansion coefficients (or generalized coordinates) is also estimated with FCNNs. In [63], the  
220 authors determine efficient approximations of the trajectories belonging to a single quantity of interest in the time domain, by mapping the parameter domain to the POD basis coefficients with artificial neural networks. The same methodology is developed in [64], where multi-layer perceptrons are employed to accurately approximate the coefficients of the reduced model for parametrized  
225 steady-state partial differential equations (PDEs). Thanks to their capability of spanning low-dimensional nonlinear manifolds, kernel-PCA have been used

in place of the more classical POD to project the high-fidelity snapshots and retain the coordinates for the following mapping task using Neural Networks [65] or Kriging-based regression [15]. The authors in [4] show that nonlinear  
230 dimensionality-reduction-based surrogate models can reduce surrogate error in sufficiently nonlinear data spaces compared to PCA-based models.

Finally to sum up, convex nonlinear dimensionality reduction techniques [35] could be applied in our application, as they are capable to capture the non-linearities of the data and potentially more efficient than linear methods as  
235 PCA. Indeed, they can allow a reduction of the number of modes that must be retained to reach a given accuracy in the reduced approximation [65]. However autoencoders appear more efficient for structured data (e.g. images, sequences) as we can incorporate prior knowledge of the data into the model by choosing the appropriate nature of the layer: for example, convolutional and recurrent layers  
240 are efficient in learning complex patterns from images and sequential data respectively. Secondly, the deep neural network architecture allows autoencoder to learn complex non-linearities with less data points compared to the aforementioned convex nonlinear dimensionality reduction [66, 67]. These reasons motivated the choice of autoencoders as dimensionality reduction technique.

### 245 2.2.2. AE-based surrogate models

Recently, models based on deep neural networks have demonstrated an outstanding performance on learning the dynamics from lower-dimensional representations. Satisfactory results have been accomplished with Convolutional Autoencoder (CAE) in the framework of emulating the spatio-temporal field  
250 governed by nonlinear PDEs [21, 22]. Here, the CAE can be considered as a nonlinear generalization of POD, with the aim at finding a reduced space to learn the dynamics from. For example, different combinations of a CAE with other machine learning algorithms were used to show the capability of AE to provide a low-dimensional model for probabilistic and deterministic predictions  
255 [37]. Similarly, CAE successfully mapped high-dimensional flow fields into a low-dimensional latent space [21, 22, 23].

An explicit relationship between input parameters and the solution over time of spatially distributed parametrized time-dependent PDEs is considered in [20, 24, 25]. The authors construct the reduced nonlinear manifold through  
260 a CAE and model the reduced dynamics (mapping input parameters to latent variables) by employing Artificial Neural Networks (ANNs).

These studies have been applied directly on structured data. However, practical experimental measurements or numerical simulations often rely on unstructured grids or non uniform sensor placements. One of the biggest challenge in this  
265 framework is that most data driven models require a uniform and regularly arranged feature matrix for each data sample, and thus their use becomes challenging for this kind of data. Several techniques [68, 69, 70, 71] have been studied to mitigate this issue, enabling the use of classic CNN also for autoencoder-based dimensionality reduction strategy when dealing with 'unstructured' data.  
270 The described works in this section relies on the application of CAE: convolutional layers have been shown to be effective for extracting representative spatial features and thus tailored for applications where dynamical systems are characterized by a state that can be represented as spatially distributed data (strong analogy with images). However, CNNs capture only the spatial dependencies  
275 which are present within the fixed-size window. Thanks to their flow control gates-based recurrent architecture (see Sect.2.3), LSTMs are coherently better to model a dynamical system and skilful at capturing short and long-term non-linear temporal correlations.

### *2.3. Notions of LSTM & LSTM-Autoencoder*

#### *280 2.3.1. Long Short-Term Memory Neural Network*

Long Short-Term Memory network is a particular variation of recurrent neural network, proposed to solve the stability gradient issues of traditional RNN [72] in handling long-range temporal dependencies. LSTM network is capable of learning linear/nonlinear temporal dependencies within sequential data, by using  
285 a recurrent mechanism and gates-based architecture. A deep LSTM neural network is composed by multiple LSTM layers which contain a series of LSTM

cells. As illustrated in Fig.2, a LSTM cell consists of 4 interacting elements: the cell state and three gates, namely the forget gate, the input gate and the output gate which control the flow of information used to update the cell state. The input gate controls the flow of input information into the cell state, whereas the forget gate regularizes the amount of information of the last cell state to be forgotten; their interaction updates the new memory of the cell state. Finally the output gate determines a part of the updated cell state to be propagated to the network.

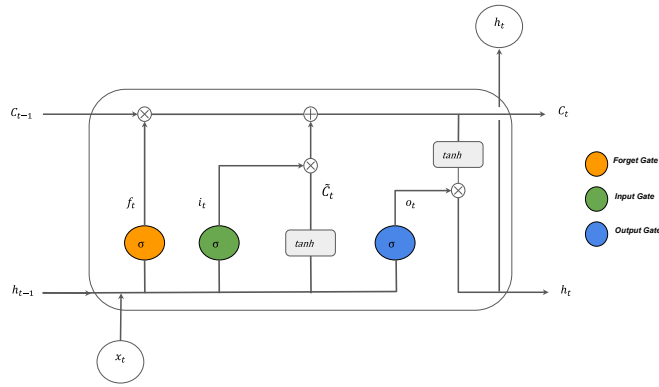


Figure 2: LSTM cell architecture of the  $j$ -th layer (index  $j$  blurred not to overwhelm the figure).

The  $j$ th layer of the LSTM neural network receives as input a sequence of vectors  $\mathbf{x}^j = \{\mathbf{x}_1^j, \mathbf{x}_2^j, \dots, \mathbf{x}_t^j, \dots\}$  where  $\mathbf{x}_t^j \in \mathbb{R}^m$  represents a  $m$ -dimensional vector at time-instance  $t$  with  $m$  equals to the output dimension of the previous layer. Denoting within the  $j$ th layer of the LSTM network, the input state to the LSTM cell as  $\mathbf{x}_t^j$ , the forget gate as  $\mathbf{f}_t^j$ , the input gate as  $\mathbf{i}_t^j$ , the output gate as  $\mathbf{o}_t^j$ , the cell state memory as  $\mathbf{C}_t^j$ , and the hidden state output as  $\mathbf{h}_t^j$  at the time-step  $t$ , the forward pass equation of an LSTM cell can be described as follows:

$$\mathbf{f}_t^j = \sigma(\mathbf{W}_f^j \cdot [\mathbf{h}_{t-1}^j, \mathbf{x}_t^j] + \mathbf{b}_f^j) \quad (1)$$

$$\mathbf{i}_t^j = \sigma(\mathbf{W}_i^j \cdot [\mathbf{h}_{t-1}^j, \mathbf{x}_t^j] + \mathbf{b}_i^j) \quad (2)$$

$$\tilde{\mathbf{C}}_t^j = \tanh(\mathbf{W}_c^j [\mathbf{h}_{t-1}^j, \mathbf{x}_t^j] + \mathbf{b}_c^j) \quad (3)$$

$$\mathbf{o}_t^j = \sigma(\mathbf{W}_o^j \cdot [\mathbf{h}_{t-1}^j, \mathbf{x}_t^j] + \mathbf{b}_o^j) \quad (4)$$

$$\mathbf{C}_t^j = \mathbf{f}_t^j \odot \mathbf{C}_{t-1}^j + \mathbf{i}_t^j \odot \tilde{\mathbf{C}}_t^j \quad (5)$$

$$\mathbf{h}_t^j = \mathbf{o}_t^j \odot \tanh(\mathbf{C}_t^j) \quad (6)$$

where the different  $\mathbf{W}_-^j$  denotes the weight matrix corresponding to the inputs of the different gates,  $\mathbf{b}_-^j$  represents the corresponding bias vectors,  $\mathbf{C}_t^j$  denotes a vector of candidate values for the cell state,  $\sigma$  is the sigmoid function,  $\tanh$  is the hyperbolic tangent function and  $\odot$  the Hadamard product. This architecture permits knowledge persistence over subsequent time-steps, retaining both long and short-term temporal patterns. Thus, the dynamical temporal behaviour of the LSTM cell makes it suitable for modelling sequential data, as shown in financial market applications [73], audio detection [74] and machine diagnosis [75, 76]. The capacity of learning sequential, non-linear time-varying patterns motivates the employment of deep LSTM neural network to model complex time-varying systems. The outcomes in [77, 78, 79, 80] shows that LSTM network is a promising, reliable approach for modelling physical dynamical systems.

### 2.3.2. LSTM-Autoencoder

The autoencoder is a (deep) neural network architecture capable of discovering relevant features of the input data in order to deliver a compressed representation. High-dimensional data can be converted to low-dimensional codes by training a bottleneck encoding-decoding multilayer neural network to reconstruct the input data, working better than PCA as a tool to reduce dimensionality [81]. Briefly, the encoder learns a compressed or latent representation of the input, whereas the decoder tries to reconstruct the input data from the

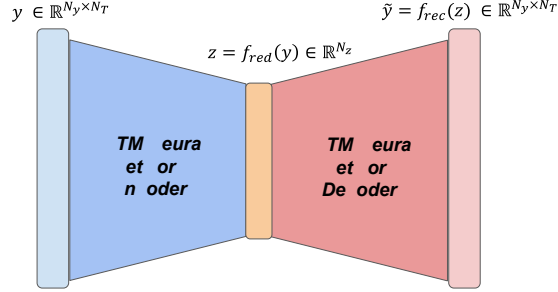


Figure 3: An illustration of a LSTM-AE network.

compressed representation. LSTM-autoencoder (Fig.3) has been successfully  
 325 implemented to extract features from multivariate time-series [82, 83]. This  
 architecture relies on the fact that LSTM network is more capable to learn im-  
 plicit temporal patterns over sequential data compared to feed-forward neural  
 networks. Auto-Encoders are used to successfully extract important features  
 and learn the high nonlinearities of data, whereas LSTMs are effective at cap-  
 330 turing temporal dependencies in sequential data. This ability makes it suitable  
 for modelling time-series data, for tasks as time-series forecasting [84], anomaly  
 detection [85] and sensors signal analysis [86].

Inspired by this, a LSTM-autoencoder is proposed to extract relevant spatio-  
 temporal features from high-dimensional multivariate time-series data and learn  
 335 a low-dimensional representation to build the surrogate model.

### 3. Proposed Approach

Let  $f$  denote a matrix-valued function on  $\Omega$ , a subset of the space  $\mathbb{R}^{N_x}$   
 ( $N_x \in \mathbb{N}$ ). For a data-driven estimation of  $f$ , a set of  $n$  observations  $\mathcal{S}_n :=$   
 $\{(\mathbf{x}_i, \mathbf{Y}_i = \mathbf{f}(\mathbf{x}_i)), i \in [1; n]\}$  is exploited, where  $\mathbf{x}_i = (x_1^i, \dots, x_{N_x}^i)^T$  represents  
 340 the input parameters and  $\mathbf{Y}_i = (\mathbf{y}_1^i, \dots, \mathbf{y}_{N_y}^i)^T$  is the corresponding solution of  
 the high-fidelity model, with  $\forall k \in [1, \dots, N_y]$ ,  $\mathbf{y}_k^i = (y_{k,1}^i, \dots, y_{k,j}^i, \dots, y_{k,N_T}^i)$  a  
 time-ordered sequence recorded for  $N_T$  time steps, describing the observable  
 physical parameters that evolve with time. Using  $\mathcal{S}_n$ , a surrogate model  $\hat{f}$  is



designed to approximate  $f : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y \times N_T}$ .

345 The proposed approach relies on the application of a dimensionality reduction technique on high-dimensional output  $\mathbf{Y}_i$ , where the high-dimensionality refers to the number of predictions that must be performed ( $N_y \times N_T$ ).

The accuracy of a surrogate model is based on the amount of information learned during the training concerning the relationship between the inputs and the  
350 outputs. When the system to approximate is nonlinear and characterized by high-dimensionality, a nonlinear dimensionality reduction technique is desired to unveil a simplified relationship between the inputs and the low-dimensional representation of the original outputs, hence facilitating the mapping task. This means that the governing function  $f$  is approximated by using the reduced output representation defined by the latent vector  $\mathbf{z}_i = (z_1^i, \dots, z_{N_z}^i)^T$ , with  $N_z$   
355 equals to the dimension of the lower-dimensional space.

An approach based on the dimensionality reduction capabilities of an LSTM-AE is developed to alleviate the burden of high-dimensionality and enable a easier classical surrogate modelling problem, i.e. approximating a vector-valued function through a FCNN. Autoencoders are used to successfully extract relevant  
360 features from data, reducing the dimensionality of the problem [87]. LSTMs are able to learn the temporal dependencies of a sequence of values [88]. Combining them together, the LSTM-AE extracts and encodes the temporal correlations and the intrinsic non-linearities of the multivariate time-series data in a lower-dimensional space. The conceived approach enables to learn a surrogate  
365 function  $\hat{g} : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_z}$  of the original problem  $f : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y \times N_T}$ , with  $N_z \ll (N_y \times N_T)$ .

Specifically, the process of the SM construction can be divided into two stages, namely *offline* and *online*. In the *offline* stage the model is trained upon the  
370 dataset  $\mathcal{S}_n$ , obtained by performing a "Design of Experiment" (DoE):  $n$  combinations of the design parameters are generated and given in input to the real-model which outputs the corresponding solutions. In our approach the *offline* stage consists on two successive steps as illustrated in Fig. 4. Firstly, the LSTM-AE is trained over the  $n$  solutions  $\mathbf{Y}_i$  in order to obtain their latent rep-

375 representations  $\mathbf{z}_i$  via the encoding function annotated  $f_{red}$ , and jointly training  
the re-construction function  $f_{rec}$ . Then, a FCNN is trained to learn the function  
 $\hat{g}$  which relates the input parameter space to the low-dimensional representa-  
tion of the physical space. The supervised learning is performed by using the  
pair samples  $\mathcal{U}_n := \{(\mathbf{x}_i, \mathbf{z}_i), i \in [1; n]\}$ , where the set of  $\mathbf{z}_i$  is extracted from  
380 the previous step. The following equations summarize the activities performed  
during the *offline* stage:

$$\mathbf{z}_i = f_{red}(\mathbf{y}_i) \quad f_{red} : \mathbb{R}^{N_y \times N_T} \rightarrow \mathbb{R}^{N_z} \quad (7)$$

$$\hat{\mathbf{z}}_i = \hat{g}(\mathbf{x}_i) \quad \hat{g} : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_z} \quad (8)$$

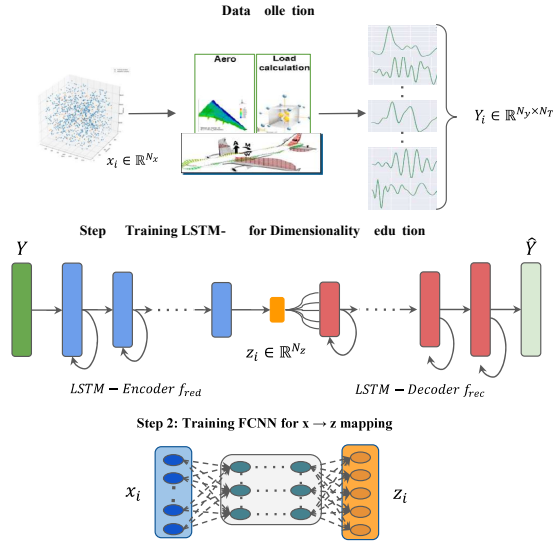


Figure 4: The "offline" process. The training phase of the proposed model consists of two networks: a LSTM-AE and a FCNN.

In the *online* stage (Fig. 5) the functions learned in the previous stage are utilized to obtain the surrogate solution corresponding to an unseen input vector, as described by Eq.(9).  
385

$$\hat{\mathbf{Y}}_i = f_{rec}(\hat{g}(\mathbf{x}_i)) \quad f_{rec} : \mathbb{R}^{N_z} \rightarrow \mathbb{R}^{N_y \times N_T} \quad (9)$$

For a new set of parameters  $\mathbf{x}_i^{new}$ , the previously trained FCNN is employed to obtain the latent representation of the solution  $\mathbf{z}_i^{new}$ . Then, the pre-trained decoder of the LSTM-AE maps the encoded representation  $\mathbf{z}_i^{new}$  into the original space, delivering the surrogate model solution  $\hat{\mathbf{Y}}_i^{new}$  associated to the input  $\mathbf{x}_i^{new}$ . In essence, the methodology enables the prediction of a high-dimensional

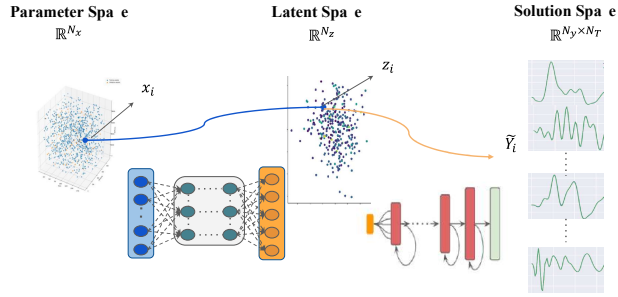


Figure 5: The "online" process. For a new point in the design phase we use the previously trained neural network to emulate the simulator output.

390

multivariate time-series initially conditioned by an exogenous vector of constant real-valued parameters. By exploiting the capabilities of LSTM-AE to capture and encode the relevant dynamical features of a multivariate time-series into lower-dimensional manifold, the high-dimensional original learning problem is transformed into a smaller and easier task, enhancing better interpolation and generalization capabilities.

395

#### 4. Empirical Study

This section presents the numerical experiments obtained by implementing our approach on two distinct parameter-dependent dynamical systems. Firstly, the applicability and the potential of our method is investigated on an academic

400

case: the LSTM-AE model is compared to a hybrid-POD (or PCA) model [63] for a simple unforced linear mass-spring-damper system. Then, the efficiency of the proposed method is assessed on the industrial aircraft landing simulations, whose dynamical behaviour is described by a high-dimensional nonlinear problem. The intuition behind the conceived method for this kind of problem is evaluated by comparing it with two main competing methodologies: hybrid-POD Neural Network [63] and 1DCAE-FCNN proposed in [25]. The experiments are performed using the Tensorflow library [89].

#### 4.1. Toy Problem: Mass-Spring-Damper System

This problem was crafted to highlight the performance benefit of using LSTM-AE-based model approach in contrast to POD-based model, when dealing with multivariate-time series dimensionality reduction. Let us consider the state space representation of the first-order ODE representing the mass-spring-damper system:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = -\frac{k}{m}x_1(t) - \frac{c}{m}x_2(t) \\ x_1(0) = x_0 \quad x_2(0) = v_0 \end{cases} \quad (10)$$

The system parameters, namely the body mass  $m$ , the damping  $c$ , and the stiffness  $k$ , govern the dynamical response of the system, in terms of frequency and amplitude of the sinusoidal waveform. Here, the task was to build a SM able to replicate the response over time of the displacement  $x_1(t)$  and its derivative  $x_2(t)$  for any set of system parameters inside a defined region of the design space.

##### 4.1.1. Experimental Setup

Two independent investigations were carried out for this application. Firstly, we performed a comparison analysis between our methodology and the hybrid-POD Neural Network model [63]. For this study, three different design of experiments (DoE) are realized to investigate the benefits of using deep autoencoder compared to POD reduction techniques, respectively with  $n = 1000$ ,  $n = 3000$

and  $n = 6000$  samples selected using a Latin Hypercube Sampling (LHS) [90] with space filling criterion in the parameter space defined by the following ranges:  $m := [100 - 900](kg)$ ,  $c := [0 - 900](Ns/m)$ ,  $k := [1000 - 91000](N/m)$ . The parameter ranges are chosen in order to obtain waveforms very different from one other and thus discover the learning capabilities of the surrogate model. The corresponding dynamic response has been generated by using an ODE integrator from Scipy library [91],  $N_T = 200$  time steps and step size  $\Delta t = 0.05$ . The initial conditions  $x_0, v_0$  are kept constant across the complete process. Thus, the learning problem is defined by  $N_T = 200$ ,  $N_y = 2$  and  $N_x = 3$ . Here,  $N_T = 200$  was chosen as trade-off between getting a wide range of damping dynamics and computational training cost.

In order to ensure the quality of the analysis, the size of the reduced manifold is the same for the two models. This implies that the number of POD basis vectors retained is equal to the latent space dimension  $N_z$ . Then, the input-to-latent mapping FCNN is trained for both approaches maintaining the same neural network architecture and hyperparameters for the ADAM [92] optimizer. The second investigation regards the dependence of the surrogate model performance on  $N_T$ . Iso-architecture LSTM-AE autoencoders and FCNN mapping were trained for the dataset with  $n = 3000$  samples with three different lengths of the corresponding temporal behaviour:  $N_T = 50$ ,  $N_T = 100$  and  $N_T = 200$ , respectively. The scope of this analysis consists on assessing how and if LSTM-AE based surrogate model is affected by  $N_T$  in terms of learning capabilities and training stability. Two problems were addressed: at the first place, it is wanted to find out if the length of the sequence impacts the strength of LSTM-AE in encoding useful information from the time-series and in providing correct outputs in a context where training and testing time series share the same length; secondly, the investigation is aimed at analyzing how the number of time steps selected for training influences LSTM-AE based surrogate model prediction capabilities on delivering different sequence lengths in testing phase.

#### 4.1.2. Numerical Results

In this section, the results of the experimental analysis upon the mass-spring-damper system are presented and discussed.

460 *Comparison between LSTM-AE based and POD based surrogate models.* The two methodologies were compared using a 10-fold cross-validation strategy by analysing the average score of the following error indicator:

$$\epsilon_{avg} = \frac{1}{n} \sum_{i=1}^n \frac{\|\mathbf{Y}^i - \hat{\mathbf{Y}}^i\|}{\|\mathbf{Y}^i\|} \quad (11)$$

where  $\mathbf{Y}^i$  and  $\hat{\mathbf{Y}}^i$  are respectively the solution of the  $i$ -th computer experiment obtained by the true model (ODE integrator) and the surrogate model (POD or LSTM-AE based). Fig. 6 shows the outcomes of the analysis: the three figures show the  $\epsilon_{avg}$  average score with respect to the dimension of the reduced manifold, with regard to reconstruction error and final surrogate prediction error. The SM based on LSTM-AE dimensionality reduction registers better performance than the hybrid POD-NN for all the three DoE, using the same FCNN mapping network. Firstly, the increase of the dataset size is beneficial for the LSTM-AE as it improves the reconstruction and the surrogate prediction, behaviour not observed using POD operation. Secondly, it can be also remarked that even increasing the number of the POD basis vectors to 48, thus with an higher number of weights, the score obtained by using the strategy LSTM-AE FCNN with a much smaller compression ratio  $cr = \frac{N_z}{(N_y \times N_x)}$  is not achieved. For the sake of completeness, the number of parameters in POD and the trainable weights in LSTM-AE are showed in Table 1. By checking out the black and blue lines of Fig 6 and the relationship with the entries of Table 1, it is possible to remark that POD method needs a larger amount of weights to achieve the same score than LSTM-AE. The reason for better reconstruction of LSTM-AE with the same number of weights, for  $n = 3000$  and  $n = 6000$  dataset, resides on the capability of the latent variables to contain robust information from data than that explained by the orthogonal basis obtained by the linear method. This characteristic is achieved by the complex operation in the LSTM

470  
475  
480

Table 1: POD and LSTM-AE latent space dimensions and weights the dimensionality reduction task. The number of parameters of LSTM-AE increases in the direction of manifold space dimension (MSD). The POD components retained are such that the number of weights of POD model and LSTM-AE are equivalent.

Manifold dimension and Model weights							
Number of weights	2618	2770	3274	4498	7810	12274	17890
LSTM-AE MSD	2	3	6	12	24	36	48
POD components ( $\simeq$ )	6	7	8	11	20	31	45

485 cell, both for the use of non-linear activation functions in an AE framework,  
as reported in [93, 94], and the gates-based mechanism which capture robust  
temporal patterns in data. On the contrary, reconstruction error with POD  
is lower in the case of  $n = 1000$  with a manifold dimension greater than 36:  
this could explained by the lack of enough data for the deep learning model.  
490 Nevertheless, the performance of the overall surrogate model based on LSTM  
is always better for the three cases. The reasons for this result are related to  
the following points: *(i)* the data distribution in the non-linear manifold allows  
the FCNN to find more robust patterns between input parameters and latent  
space; *(ii)* the decoder is more robust to possible errors in mapping the reduced  
495 dimension coefficients used to return to the original space, as the autoencoder  
provides better interpolation capabilities in the reduced manifold space.

Figure 6 suggests another interesting cue: one would expect a decreasing trend  
of the LSTM-AE reconstruction error in the direction of the manifold space  
dimension. However, this is not particularly true in this case as changing the  
500 manifold space size leads to a substantial modification of the number of weights.  
The  $\pm 1.96\sigma$  interval over the LSTM-AE reconstruction 10-fold cross-validated  
mean error indicates that under a certain dimension of the latent space, and thus  
below a certain number of trainable parameters, the trainings of the LSTM-AEs  
are highly noisy and quite unstable. It has been already demonstrated that the  
505 performance of an autoencoder depends on the number of tunable weights [95].

In this case, LSTM-AE appears to be very sensible to the amount of trainable parameters in terms of weights updates and in the stability towards loss convergence for the reconstruction error.

In conclusion, even for a linear system, as the one presented in Eq.(10), the hybrid POD-NN model needs a large amount of basis function to achieve a prediction score similar to that we remark for the AE-based model. This result leads to assume that the number of coefficients defining the linear reduced manifold would be considerably larger in relation to those of the nonlinear manifold for building the reduced-order model. This could represent a significant problem for more complex systems, with higher number of input parameters and time-series outputs. Indeed, having more variables to predict, the mapping FCNN would struggle in approximating the function lying the input parameters  $\mathbf{x}_i$  to low-dimensional representation  $\mathbf{z}_i$ , consequently increasing the difficulty to reconstruct the solution in the original space.

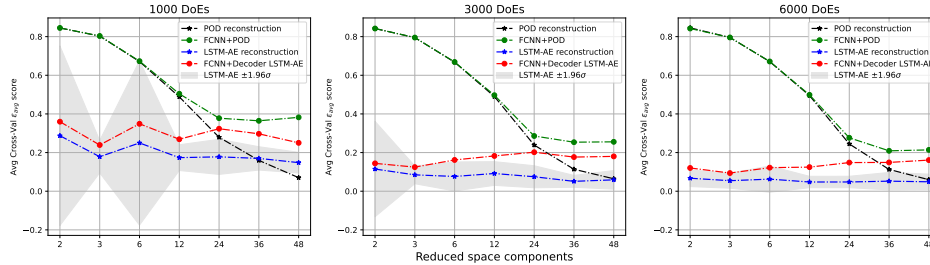


Figure 6: Cross validation mean error Indicator (y-axis) vs Dimension of the reduced manifold (x-axis) for the three design of experiments for the Mass-Spring-Damper problem.

*The influence of  $N_T$  on LSTM-AE based surrogate model performance.* The investigation about the dependence of  $N_T$  on LSTM-AE based surrogate model performance is summarized in Figure 7 in the form of dot plot. The cross grid-based evaluation allows to characterize the influence of  $N_T$  used for training when a prediction of a time series, for an equal or different number of time steps, is demanded during the test phase. In detail, each point represents the mean validation set error which derives by averaging the results from different



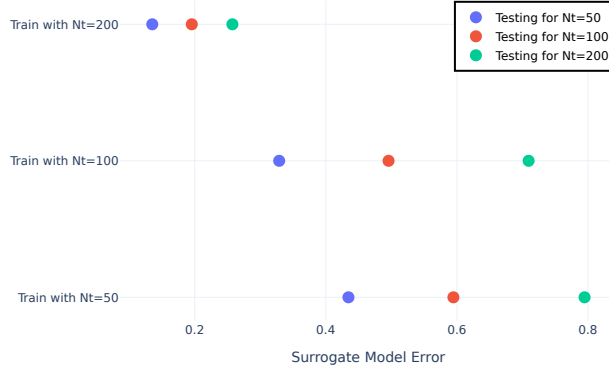


Figure 7: Dot plot showing the influence of  $N_T$  in LSTM-AE based surrogate model. For each choice of  $N_T$  for model training (y-axis), three different temporal window prediction have been analyzed. The x-axis refers to the error indicator of Eq.11.

latent space dimension models. The results show that (i) by increasing the number of time steps of the sequence in test phase, the error increases due to error propagation caused by the recursive mechanism of the prediction; (ii) at iso- $N_T$  in testing, it is preferable training the model with longer sequences as the LSTM is able to retain more information and indeed to improve prediction capability.

#### 4.2. Industrial Case: Dynamic Landing Loads Simulation

During landing phase, the aircraft is exposed to a short-duration impulsive impact that results in a dynamic response of the airframe and of the landing gear. The observable physical quantities, arising from the interaction between the aircraft and the ground, represent forces, moments and accelerations evolving over time (Interesting Quantities -IQs). Critical IQs largely determine the size of both the airframe and landing gear structures. This means that the loads arising at touchdown have to be determined accurately for the structure to be adequately dimensioned. The adequacy of large transport category airplane to

the flight is regulated by the airworthiness authorities (EASA CS 25 and FAR 25). The aircraft is required to comply with the directives for a set of design conditions. Some of the design conditions require to take into account the dynamics of the aircraft. For the dynamic landing load conditions (CS 25.473, CS 25.479, CS 25.481), a full time history of the touch down has to be computed. A complex process is set-up to represent the static and possibly dynamic response of the aircraft to compute and validate dynamic landing loads. In this framework, the aircraft is subjected to applied external loads, accounting also for distributed inertia, aerodynamic and structural forces and moments. The ground loads model for dynamic simulations consists of two coupled models: a flexible aircraft, defined by a Finite Element Model (FEM) and a Mass Model modal representation; and the landing-gear model, defined by its kinematic, the inertial model and the dynamic flexibility (in terms of normal modes). The overall model presents high non-linearities which arise from the landing gear model. The equations describing the response of the shock absorber introduce nonlinear terms: exponential stroke-force relation, square-dependent damping force, internal friction (step functions) and stroke limiters are source of high non-linearities and a linearisation is not possible. Here the dynamic simulation is carried out using a symmetric aircraft model with a symmetric configuration, without any control law system implemented. Such a configuration, in addition to arbitrary CS25 requirements that aim to cover all possible asymmetry coming either from the landing conditions or from the aircraft itself, satisfies the requirements for airframe and landing gear structural design. Nonetheless, it is worth drawing attention also to aircraft landing operational service where the aircraft is asymmetrically loaded, either coming from inherent properties of the airplane, or external conditions resulting in roll, slideslip and yaw rates or as a consequence of a system failure. Indeed, in the absence of a sophisticated fly-by-wire control systems, airplanes with non-aerodynamic asymmetric loads require a safe flight boundary analysis on the basis of the available lateral flight control authority [96]. This issue has been pointed out by the authors in [97, 98] who based themselves on general rotational analysis [99] to define a methodology to determine

the supplement to the safe flight boundaries of symmetrically loaded airplanes for asymmetric load situations within terminal flight phases. Such situations are investigated with rational calculations for in-service event analysis but are not covered by the dynamic landing simulator used here. The aircraft landing loads depend on a set of design parameters, related both to landing gear and aircraft models. Exploring the neighbourhood of a defined design, i.e. investigating how the  $IQs$  are impacted by the multivariate percentage variations of design parameters, is time-consuming with traditional process (construction and assemblage of mass model, airframe and landing gear FEMs; Computer Aided Engineering (CAE) Nastran-based simulations and post-processing). This leads to the need for a Surrogate Model which allows a fast exploration of the parameters design space, maintaining a satisfying level of accuracy. A data-driven model can capture non-linear relationships of varying multiple inputs at the same time rather than performing an investigation on the effects of varying the parameters one at a time with the high-fidelity process. In this analysis, the inputs of the overall simulation process to emulate via our SM (see Fig.8) is represented by a set of constant values containing information about the two aforementioned submodels, i.e. aircraft flight state, the landing gear model and the mass distribution along the fuselage and the wings. The simulation outputs the  $IQs$  for all the sections of the different aircraft components. The goal is to predict the  $IQs$  for new (i.e. unseen during training phase) landing gear configurations and fuselage mass distributions combined with unchanged critical flight cases.

It is worth highlighting two crucial points when predicting the multivariate time-series representing the aircraft response. The sizing of the aircraft is extremely dependent on the maximum loads computed during the simulation: the so-called 1D envelope is calculated by finding the maximum (and minimum) peaks over all the time series and configurations. However, it is not sufficient to compute these values because it is essential to take into account the temporal correlations between different  $IQs$  - *Correlated Loads* - as the principal stresses depend upon 2D or 3D loads. Thus, the SM should be able to accurately predict the maximum peaks and the temporal dynamic response of all the  $IQs$ .

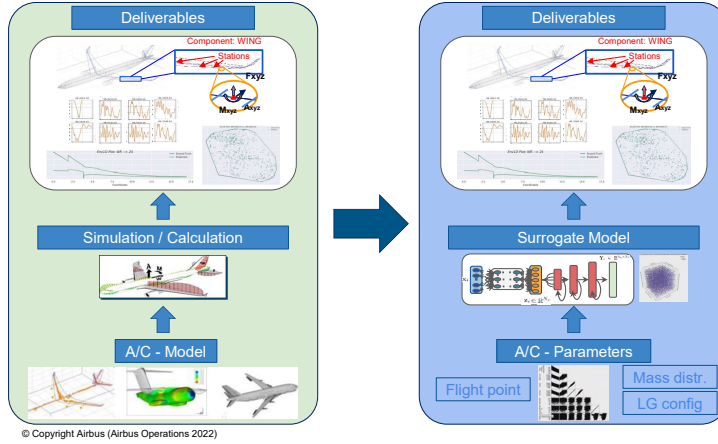


Figure 8: From a high-fidelity computationally expensive process to an agile and flexible process enabled by surrogate modelling.

#### 4.2.1. Experimental Setup

605 The input of our model is a real-valued vector  $\mathbf{x}_i$ , composed by  $N_x = 26$  design parameters. These parameters refer to the pitching angle, the dynamic pressure, the aerodynamic center, the shock absorber characteristics of the landing gear, the center of gravity, the momenta of inertia for the three axes and finally 3 parameters define the mass distribution. Thus, each  $\mathbf{x}_i$  defines a different load case. For each load case, the simulation model is run and automatically solved: the outcome is the dynamic aircraft response described by a multivariate time-series composed by  $N_y = 711 IQs$  with  $N_T = 101$  time steps. Here,  $N_T = 101$  corresponds to 1 second of simulation: it is within this time-frame that the dynamics of the landing takes place and the peaks appear, according to the design parameters. As the objective is to learn and predict, besides the dynamics, also where the peaks in the temporal axis occur within this temporal window, modifying  $N_T$  would conceal information during the training and would not allow to spot the peaks in testing. For these reasons  $N_T = 101$  has been fixed for both training and testing. The dataset is composed by 42282

615

620 pairs  $(\mathbf{x}_i, \mathbf{Y}_i)$  and was split into 60% for training, 20% for validation and 20%  
 for the final test. The split has been performed randomly, ensuring that the  
 statistical distributions for the three datasets are consistent within each other.  
 The validation set is used to provide an unbiased evaluation of a model fit on  
 the training dataset while tuning model hyperparameters; as a result of this  
 625 phase, the final model is retained, and its performance is assessed in the test  
 set by means of defined metrics. The load cases to build the dataset are set  
 up as follows: 38 critical flight points are selected, according to the most sig-  
 nificant landing scenarios on the aircraft design loads. These critical points are  
 retained and combined with several combinations of landing-gear parameters  
 630 and mass-fuselage distributions. A Latin Hypercube Sampling (LHS) strategy  
 was adopted for designing 1001 combinations of the 8 landing gear parameters,  
 obtained as percentage variation of the given reference values. Then, the mass  
 fuselage distributions for the 20 most critical mass cases (out of 38) are varied  
 with a randomized shape function, under the condition to maintain the total  
 635 mass and the center of gravity constant. The resulting 8916 new mass distri-  
 butions are randomly combined with the above 1001 landing gear combinations  
 and the critical flight cases to generate the dataset.

For the training task, a standard scaling is performed for the input parameters  
 data  $\mathcal{S}_n^x = \{\mathbf{x}_i, i \in [1; n]\}$ . A standard scaling is applied also the solution data  
 640  $\mathcal{S}_n^y = \{\mathbf{Y}_i, i \in [1; n]\}$ , where each  $IQ \mathbf{y}_k^i$  is scaled using the mean and the stan-  
 dard deviation computed over all load cases and over all the time step for all  
 the  $IQs$ .

The architecture of our LSTM-AE and the learning parameters are selected as  
 the result of a grid search procedure. The same strategy is adopted to select  
 645 the best FCNN configuration. Table 2 describe the final LSTM-AE and FCNN  
 architectures retained. Both networks are trained for 5000 epochs with batch  
 size equal to 128 and using the ADAM optimizer with learning rate equal to  
 $1e^{-3}$ , to minimize the mean squared error (MSE) loss objective.

Moreover, an early stopping strategy [100] is employed in both trainings to avoid  
 650 overfitting. Several LSTM-Autoencoders are trained with different latent space

Table 2: LSTM-AE & FCNN Neural Networks Architecture with  $N_T = 101$ ,  $N_y = 711$ ,  $N_x = 26$ ,  $N_z = 150$  and  $n$  equals to the batch size.

Component	Encoder	Decoder	FCNN
<b>Layers</b>	Input( $n, N_T, N_y$ )	RepeatVector( $N_T$ )	Input( $n, N_x$ )
	LSTM( $n, N_T, 640$ )	LSTM( $n, N_T, 256$ )	Dense( $n, 256$ )
	LSTM( $n, N_T, 256$ )	LSTM( $n, N_T, 640$ )	Dense( $n, 256$ )
	LSTM( $n, 1, N_z$ )	TimeDistr. ( $n, N_T, N_y$ )	Dense( $n, 256$ )
			Dense( $n, N_z$ )

dimension, maintaining the same encoder and decoder structures, to find the architecture yielding the best SM accuracy. Fig. 9 illustrates the RMSE validation error behaviour with respect to the dimension of the latent space, compared to POD dimensionality reduction method. Also in this case, the error does not decrease increasing the dimension of the latent space. However, in contrast to the previous application, the zig-zag like behaviour does not appear because the relative 'delta' of weights between the models is small, given the considerable number of LSTM-AE trainable weights. The minimum reconstruction error for the LSTM-AE and FCNN mapping, and for the overall SM model are obtained for a latent space dimension equal to 150. The valley like behaviour confirms the influence of the amount of model weights on the learning process: the model with latent space dimension equals to 150 represents the optimal network in terms of weight optimization, achieving the lowest reconstruction error. The results show that even if POD reconstruction is better than LSTM-AE for manifold size equals to 250, the overall surrogate performance using the same FCNN is lower: the reason for such behaviour is related to what has been argued in Sec. 4.1.2, i.e. the activation functions mechanism of the LSTM cell allows to encode robust information, consequently ensuring strong reconstruction capabilities and enhancing the FCNN mapping. This point is emphasized by analyzing the performance of LSTM-AE and POD at iso number of weights. The different manifold sized LSTM-AE models have the same number of weights than POD model when for the latter the number of components retained is within

the range [104 – 114]: comparing the errors of the entire blue line and the black line in that range, a better performance in reconstruction of LSTM-AE can be remarked.

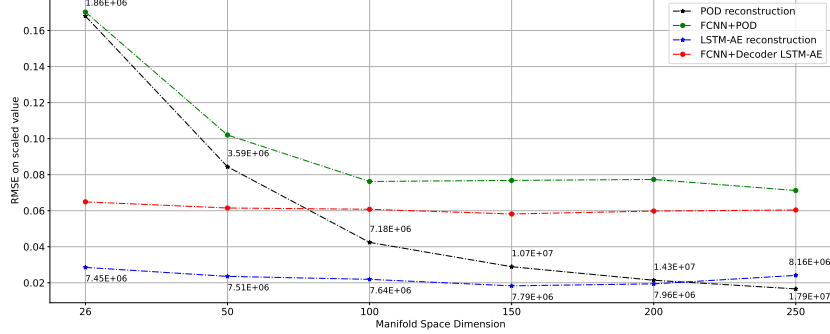


Figure 9: RMSE vs Dimension of latent space for POD and LSTM-AE methods. The values refer to the number of parameters for the dimensionality reduction model.

675

#### 4.2.2. Numerical Results

The metric in Eq.(11) is adopted to evaluate the performance of LSTM-AE. Moreover, as explained at the end of Sec.4.2, the maximum and minimum peaks of each IQ take on great importance for the preliminary design of 1D and 2D envelopes: the model should accurately predict not only the peak values, but mainly predict them in the exact time step they take place in order to take into account the correlated loads [101] prediction. For this, the following expressions are employed as additive metrics to assess the capability of the model :

680

$$MAPE_+ = 100 \times \frac{1}{nN_y} \sum_{i=1}^n \sum_{j=1}^{N_y} \left| \frac{Y_{ij}^{t=t_{max}} - \tilde{Y}_{ij}^{t=t_{max}}}{Y_{ij}^{t=t_{max}} + \alpha} \right| \quad (12)$$

$$MAPE_- = 100 \times \frac{1}{nN_y} \sum_{i=1}^n \sum_{j=1}^{N_y} \left| \frac{Y_{ij}^{t=t_{min}} - \tilde{Y}_{ij}^{t=t_{min}}}{Y_{ij}^{t=t_{min}} + \alpha} \right| \quad (13)$$

The scores measure, for each  $N_y$ -th IQ and for each  $n$ -th load validation case, the difference between the true peak value (maximum and minimum peaks de-

685

fined respectively by  $Y_{ij}^{t=t_{max}}$  and  $Y_{ij}^{t=t_{min}}$ ) and the value predicted by the SM at the time step the peak takes place. The recurrent autoencoder based surrogate model predictions are compared to the random-search optimized hybrid POD-FCNN, and to the 1D-CAE FCNN architecture. The results are summarized in Table 3. Analyzing the results, it can be remarked that extracting enough

Table 3: Scores achieved by LSTM-AE, 1D-CAE ad hybrid POD-NN SM on the industrial dataset (the lower the value, the better the SM performs). RMSE refers to the best error on the scaled validation set during the hyperparameter tuning phase. The final model is assessed using the metrics computed on test set. For each the dimensionality reduction models (LSTM-AE, POD and CNN1D-AE), the dimension of the reduced space and the number of weights are showed.

Model	Val. error	Metrics		
	RMSE	$\epsilon_{avg}$	MAPE <sub>+</sub>	MAPE <sub>-</sub>
LSTM-AE [z=150] FCNN (7.792.687 weights)	<b>0.0582</b>	<b>0.0066</b>	<b>7.72%</b>	<b>8.8%</b>
1D-CAE [z=150] FCNN (7.103.157 weights)	0.07863	0.0112	18.52%	25.11%
Hybrid PD-NN [z=150] FCNN (10.771.650 weights)	0.07698	0.0096	11.17%	16.41%
1D-CAE [z=150] FCNN (16.092.765 weights)	0.07697	0.0116	16.48%	20.15%
Hybrid POD-NN [z=250] FCNN (17.952.750 weights)	0.0712	0.0080	11.88%	14.59%

690

POD modes to preserve the energy content of the multivariate time-series data is more powerful than exploiting local-low level features through convolutional layers that lead to a nonlinear reduced manifold, despite having larger number of weights than the LSTM-AE. On the other hand, capturing and encoding the temporal correlations via LSTM cells is much more effective for this kind of problem, both for predicting the multivariate time series and for accurately estimating the peaks. Fig. 10 illustrates how accurate LSTM-AE FCNN is with respect to the other models. LSTM-AE FCNN solution is the one which better predicts the amplitude of the oscillation peaks and the behaviour of the waveform over time. In all the figures of this section the y-axis, the name of the load

700



cases and components sections are blurred due to confidentiality.

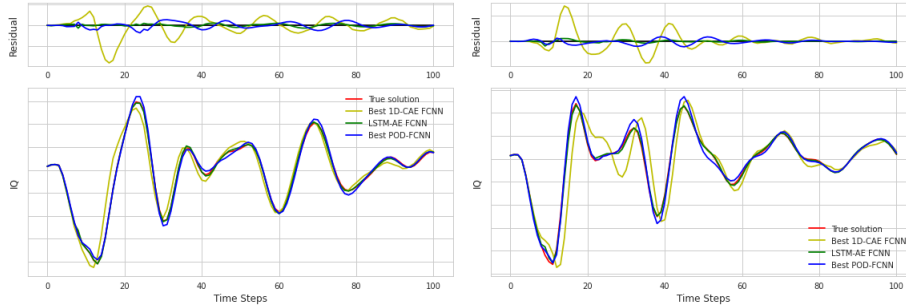
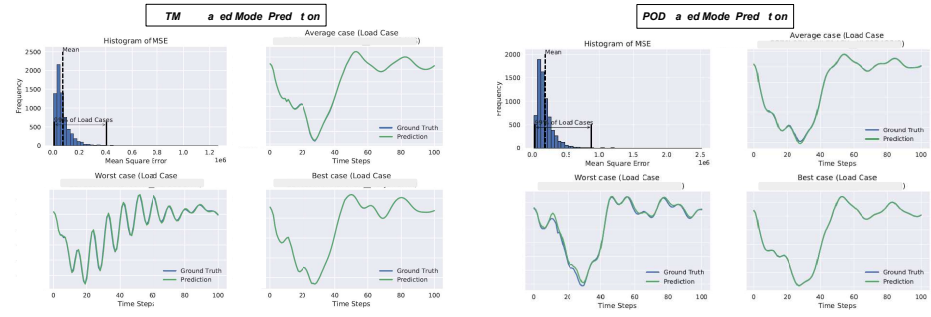


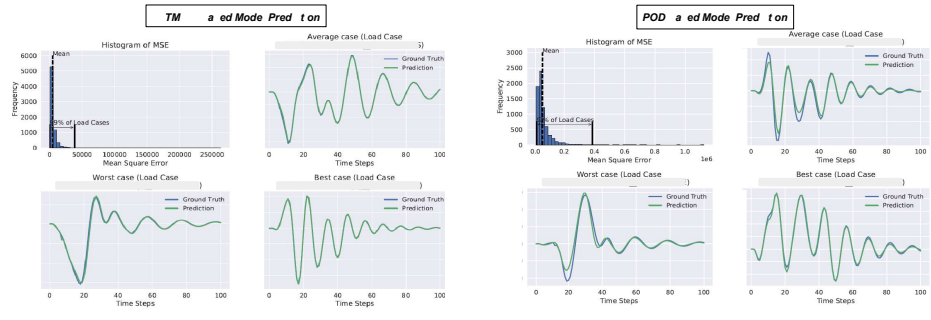
Figure 10: Comparison of the exact solution vs SMs predictions of the same IQ for two validation load cases. Prediction with LSTM-AE based surrogate model presents a lower residual error.

In the following, only hybrid-POD SM is accounted as main benchmark to assess our model, as it is more accurate than 1D-CAE based model. Fig. 11 compares the prediction of LSTM-AE and hybrid-POD SMs for three  $IQ$ s: fuselage (Fig.11a), left wing (Fig.11c) and main landing gear (Fig.11b). The MSE histograms are computed considering all the load cases not included in the training set: LSTM-AE based model produces more accurate predictions, as the right-skewed distribution statistics (mean and mode) are shifted towards lower values with respect to the PCA-based model ones; moreover it can be highlighted that the largest MSE value is much lower in the LSTM-AE based model case (in some case more than 50% error reduction). The outcomes showed for these examples are generalizable to all the other  $IQ$ s.

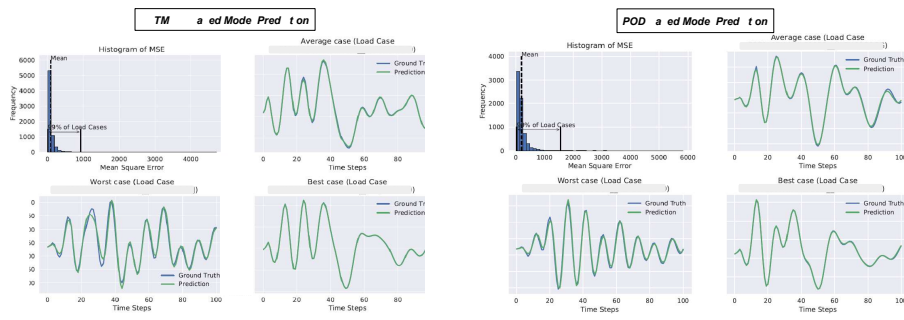
The ability of the LSTM-AE FCNN to predict both the sequence trend and to spot the maximum peaks in a more efficient way than hybrid-POD model is highlighted in Fig. 12, Fig. 14 and with the 2D envelope in Fig 13. As previously explained (see Sec.4.2), 2D envelopes have a great interest for airframe stress design. Thus, besides an accurate 1D envelope prediction, for which our SM could have been trained simply for vector-valued outputs (only the peaks of the time-series), it is necessary also consider the temporal correlation between different  $IQ$ s.



(a) Vertical Force Forward Fuselage Section



(b) Main Landing Gear Bending Moment



(c) Vertical Force Outer Wing Section

Figure 11: An example of comparison between LSTM-AE based SM vs PCA-based SM predictions: time-series for Best, Average and Worst MSE errors load cases.

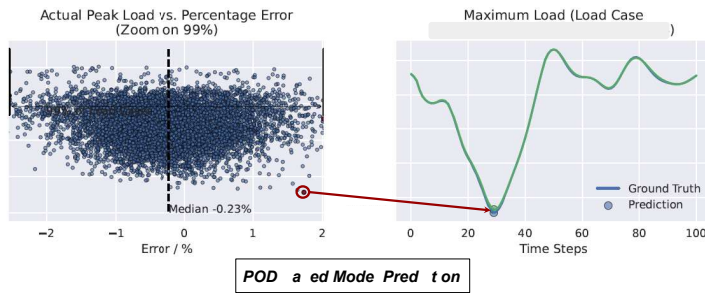
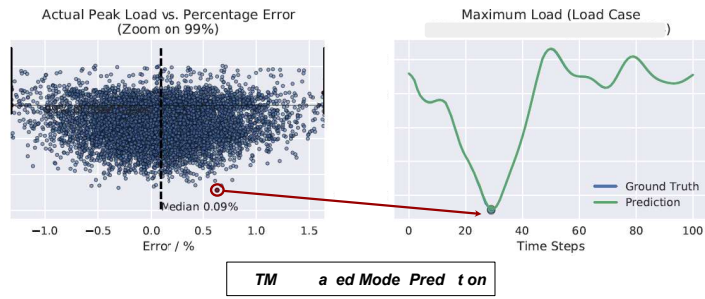


Figure 12: Peaks prediction of vertical force in one fuselage section for the 99% load cases within the validation set. The plot on the right represents the IQ time-series for the critical load case. The red point value is used to compute the 1D envelope.

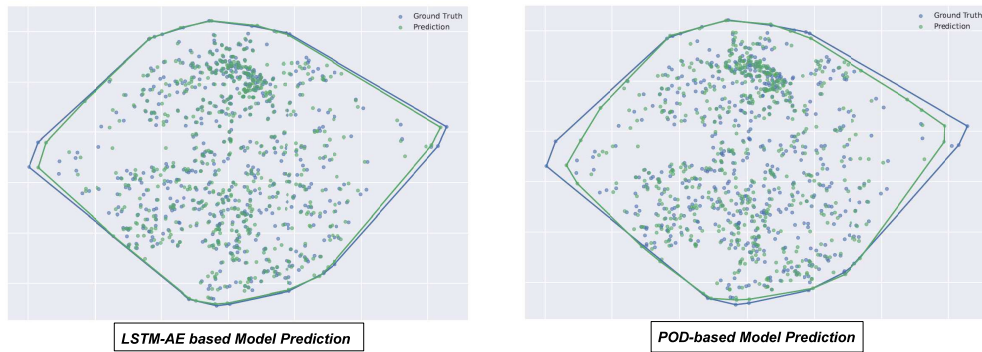


Figure 13: Hybrid-POD FCNN vs LSTM-AE FCNN in predicting 2D envelope. 2D envelope for lateral vs vertical forces for the same fuselage station; blue dots represent the truth values, the green ones are predicted by the SM. The solid blue and green lines define the envelope along the complete validation dataset

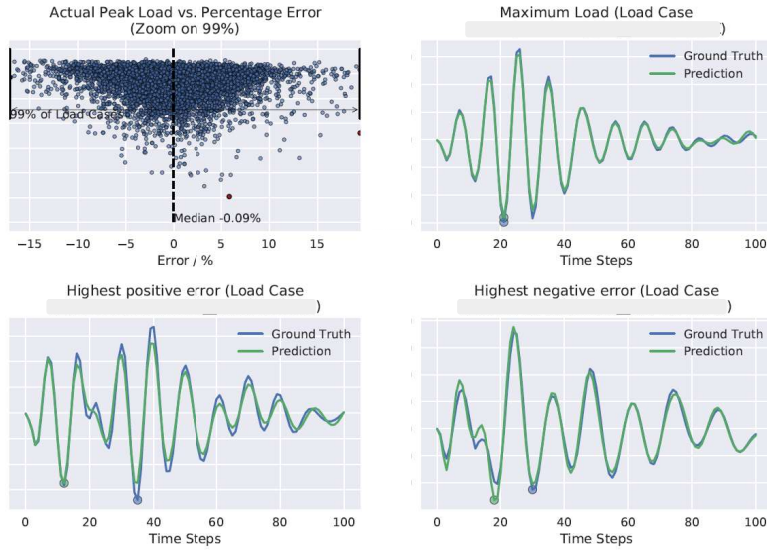


Figure 14: Hybrid-POD FCNN larger peak errors and discrepancy in time lead to inaccurate 2D envelope.

Indeed, the accuracy of a 2D envelope prediction is not only dependent by the peak values but also by the temporal behaviour of the two  $IQ$ s. Fig. 12 compares the minimum peak predictions for a singular  $IQ$ . It can be observed that the critical load case (red point), i.e. the minimum value attained by the  $IQ$  over all the load cases considered, is predicted with an error less than 1% only by the LSTM-AE SM. The same result has been registered for almost all the other  $IQ$ s. This is a crucial outcome because in order for the predictions to be used by the engineers for sizing the aircraft with the so-called 1D envelope, the error cannot exceed 1% for the most critical load cases. Moreover, the inability of the SM to exactly predict the peaks deeply affects the accuracy of the 2D envelope estimate (Fig. 13): the problem is not only related to the peak value itself but also at the time it occurs. In the hybrid-POD model for example, the discrepancy between blue and green solid lines is due to both the inability to

capture the temporal correlation (see Fig.14) and to predict the exact value of  
735 the peaks.

## 5. Conclusion and Future Work

An end-to-end methodology based on an LSTM-autoencoder coupled with a FCNN has been conceived for surrogate modeling of design parameter-dependent dynamic landing correlated loads to predict 1D and 2D load envelopes for structural design. The methodology has been validated by comparing it to alternative  
740 techniques on two different problems: the one degree of freedom linear mass-spring-damper system and the high-dimensional nonlinear problem representing the ground loads simulation of a flexible aircraft. The experiments show that the LSTM-AE dimensionality-reduction-based model outperforms both classical method (POD) and non-linear feed-forward networks (CNN1D-AE) when  
745 non-linear time-lag dependencies are present on data. The reasons of better performance are definitely related to the recurrent gates-based mechanism of LSTM which allows to automatically capture temporal dependencies. Indeed, even using convolutional layers coupled with non-linear activation functions, already acknowledged to be better than POD operation, with a larger number of  
750 trainable weights do not lead to comparable prediction performance than that delivered by LSTM-based AE.

The presented results on the two dynamical systems verify the applicability of the model to any task where the dynamic response conditioned on an exogenous  
755 set of constant parameters has to be predicted, mainly in a high-dimensional context. This work could be extended to efficiently quantify the effects of uncertainty in the system parameters and randomness in network modeling. It could be possible to combine the deep learning model with bayesian inference (Bayesian Neural Network - BNN) to estimate the uncertainty related to the  
760 network prediction [102], critical in this kind of context, where errors could lead to an over/under design of structure. As well as for BNN, it would be interesting to investigate variational inference within the autoencoder framework (VAE)

[103] to improve generative capabilities. Other suggestions for future work relies on improving the generalization ability of the algorithm via physics embedding [104], improving interpolation in autoencoders [105], exploring alternative data encoding methods to estimate the intrinsic dimension of a complex system from data and on loss function modification to better capture the presence of critical load peaks.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgment

This work has been supported by Airbus Operations SAS and the French National Agency for Technological Research (ANRT) within the CIFRE framework (grant N° 2019/1815).

### References

- [1] A. I. J. Forrester, A. Sóbester, A. J. Keane, *Engineering Design via Surrogate Modelling*, John Wiley & Sons, Ltd, 2008, Ch. 2, pp. 33–76. doi:<https://doi.org/10.1002/9780470770801.ch2>.
- [2] J. Tao, G. Sun, Application of deep learning based multi-fidelity surrogate model to robust aerodynamic design optimization, *Aerospace Science and Technology* 92 (2019) 722–737. doi:<https://doi.org/10.1016/j.ast.2019.07.002>.
- [3] M. Rippepi, M. Verveld, N. Karcher, et al., Reduced-order models for aerodynamic applications, loads and mdo, *CEAS Aeronaut Journal* 9 (2018) 171–193. doi:<https://doi.org/10.1007/s13272-018-0283-6>.
- [4] G. D. Bird, S. E. Gorrell, J. L. Salmon, Dimensionality-reduction-based surrogate models for real-time design space exploration of a jet engine compressor blade, *Aerospace Science and Technology* 118 (2021) 107077. doi:<https://doi.org/10.1016/j.ast.2021.107077>.
- [5] X. Bertrand, F. Tost, S. Champagneux, Wing airfoil pressure calibration with deep learning, *AIAA Aviation Forum* (2019). doi:[10.2514/6.2019-3066](https://doi.org/10.2514/6.2019-3066).

- [6] R. K. Tripathy, I. Bilonis, Deep UQ: Learning deep neural network surrogate models for high dimensional uncertainty quantification, *Journal of Computational Physics* 375 (2018) 565–588. doi:<https://doi.org/10.1016/j.jcp.2018.08.036>.
- 795 [7] R. Dupuis, J.-C. Jouhaud, P. Sagaut, Surrogate modeling of aerodynamic simulations for multiple operating conditions using machine learning, *AIAA Journal* 56 (9) (2018) 3622–3635. doi:[10.2514/1.J056405](https://doi.org/10.2514/1.J056405).
- [8] H. Ganti, P. Khare, Data-driven surrogate modeling of multiphase flows using machine learning techniques, *Computers & Fluids* 211 (2020) 104626. doi:<https://doi.org/10.1016/j.compfluid.2020.104626>.
- 800 [9] K. Li, J. Kou, W. Zhang, Unsteady aerodynamic reduced-order modeling based on machine learning across multiple airfoils, *Aerospace Science and Technology* 119 (2021) 107173. doi:<https://doi.org/10.1016/j.ast.2021.107173>.
- [10] W. Li, X. Gao, H. Liu, Efficient prediction of transonic flutter boundaries for varying mach number and angle of attack via lstm network, *Aerospace Science and Technology* 110 (2021) 106451. doi:<https://doi.org/10.1016/j.ast.2020.106451>.
- 805 [11] S. L. Brunton, J. Nathan Kutz, K. Manohar, A. Y. Aravkin, K. Morgansen, J. Klemisch, N. Goebel, J. Buttrick, J. Poskin, A. W. Blom-Schieber, T. Hogan, D. McDonald, Data-driven aerospace engineering: Reframing the industry with machine learning, *AIAA Journal* 59 (8) (2021) 2820–2847. doi:[10.2514/1.J060131](https://doi.org/10.2514/1.J060131).
- 810 [12] J. Biannic, G. Hardier, C. Roos, C. Seren, L. Verdier, Surrogate Models for Aircraft Flight Control: Some Off-Line and Embedded Applications, *Aerospace Lab* (12) (2016) pages 1–21. doi:[10.12762/2016.AL12-14](https://doi.org/10.12762/2016.AL12-14).
- [13] J. T. Hwang, J. R. Martins, A fast-prediction surrogate model for large datasets, *Aerospace Science and Technology* 75 (2018) 74–87. doi:<https://doi.org/10.1016/j.ast.2017.12.030>.
- 815 [14] Y. Zhou, Z. Lu, J. Hu, Y. Hu, Surrogate modeling of high-dimensional problems via data-driven polynomial chaos expansions and sparse partial least square, *Computer Methods in Applied Mechanics and Engineering* 364 (2020) 112906. doi:<https://doi.org/10.1016/j.cma.2020.112906>.
- [15] C. Lataniotis, S. Marelli, B. Sudret, Extending classical surrogate modelling to high dimensions through supervised dimensionality reduction: a data-driven approach, *International Journal for Uncertainty Quantification* 10 (01 2020). doi:[10.1615/Int.J.UncertaintyQuantification.2020031935](https://doi.org/10.1615/Int.J.UncertaintyQuantification.2020031935).
- 820 [16] J. Morton, M. J. Kochenderfer, F. D. Witherden, Parameter-conditioned sequential generative modeling of fluid flows, *AIAA Journal* 59 (3) (2021) 825–841. doi:[10.2514/1.J059315](https://doi.org/10.2514/1.J059315).
- 825 [17] Z. Wang, S. Rosa, B. Yang, S. Wang, N. Trigoni, A. Markham, 3D-Physnet: Learning the intuitive physics of non-rigid object deformations, in: *Proceedings of the 27th International Joint Conference on Artificial Intelligence, IJCAI'18*, AAAI Press, 2018, p. 49584964. doi:<https://doi.org/10.24963/ijcai.2018/688>.

- 830 [18] R. Maulik, K. Fukami, N. Ramachandra, K. Fukagata, K. Taira, Probabilistic neural networks for fluid flow surrogate modeling and data recovery, *Phys. Rev. Fluids* 5 (2020) 104401. doi:10.1103/PhysRevFluids.5.104401.
- [19] K. Lee, E. J. Parish, Parameterized neural ordinary differential equations: applications to computational physics problems, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 477 (2253) (2021) 20210162. doi:10.1098/rspa.2021.0162.
- 835 [20] S. Fresca, L. Dede', A. Manzoni, A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs, *Journal of Scientific Computing* 87 (2021). doi:https://doi.org/10.1007/s10915-021-01462-7.
- [21] R. Maulik, B. Lusch, P. Balaprakash, Reduced-order modeling of advection-dominated systems with recurrent neural networks and convolutional autoencoders, *Physics of Fluids* 33 (3) 840 (2021) 037106. doi:10.1063/5.0039986.
- [22] T. Nakamura, K. Fukami, K. Hasegawa, Y. Nabaie, K. Fukagata, Convolutional neural network and long short-term memory based reduced order surrogate for minimal turbulent channel flow, *Physics of Fluids* 33 (2) (2021) 025116. doi:10.1063/5.0039845.
- [23] F. J. Gonzalez, M. Balajewicz, Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems, *ArXiv abs/1808.01346* (2018).
- 845 [24] N. R. Franco, A. Manzoni, P. Zunino, A deep learning approach to reduced order modelling of parameter dependent partial differential equations, *CoRR abs/2103.06183* (2021).
- [25] S. Nikolopoulos, I. Kalogeris, V. Papadopoulos, Non-intrusive surrogate modeling for parametrized time-dependent partial differential equations using convolutional autoencoders, *Engineering Applications of Artificial Intelligence* 109 (2022) 104652. doi:https://doi.org/10.1016/j.engappai.2021.104652.
- 850 [26] M. J. Asher, B. F. W. Croke, A. J. Jakeman, L. J. M. Peeters, A review of surrogate models and their application to groundwater modeling, *Water Resources Research* 51 (8) (2015) 5957–5973. doi:https://doi.org/10.1002/2015WR016967.
- 855 [27] M. Verleysen, D. François, The curse of dimensionality in data mining and time series prediction, in: J. Cabestany, A. Prieto, F. Sandoval (Eds.), *Computational Intelligence and Bioinspired Systems*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 758–770. doi:https://doi.org/10.1007/11494669\_93.
- [28] I. Jolliffe, Principal component analysis, in: M. Lovric (Ed.), *International Encyclopedia of Statistical Science*, Springer Berlin Heidelberg, 2011, pp. 1094–1096. doi:10.1007/978-3-642-04898-2\_455.
- 860 [29] M. Rathinam, L. R. Petzold, A new look at proper orthogonal decomposition, *SIAM Journal on Numerical Analysis* 41 (5) (2003) 1893–1925. doi:10.1137/S0036142901389049.
- [30] T. Zhou, Y. Peng, Kernel principal component analysis-based gaussian process regression modelling for high-dimensional reliability analysis, *Computers & Structures* 241 (2020) 106358. doi:https://doi.org/10.1016/j.compstruc.2020.106358.
- 865



- 870 [31] B. Schölkopf, A. Smola, K.-R. Müller, Kernel principal component analysis, in: W. Gerstner, A. Germond, M. Hasler, J.-D. Nicoud (Eds.), *Artificial Neural Networks — ICANN'97*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997, pp. 583–588. doi:<https://doi.org/10.1007/BFb0020217>.
- [32] J. B. Tenenbaum, V. de Silva, J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, *Science* 290 (5500) (2000) 2319–2323. doi:[10.1126/science.290.5500.2319](https://doi.org/10.1126/science.290.5500.2319).
- 875 [33] R. R. Coifman, S. Lafon, Diffusion maps, *Applied and Computational Harmonic Analysis* 21 (1) (2006) 5–30, special Issue: Diffusion Maps and Wavelets. doi:<https://doi.org/10.1016/j.acha.2006.04.006>.
- [34] S. Ullah, D. A. Nguyen, H. Wang, S. Menzel, B. Sendhoff, T. Bck, Exploring dimensionality reduction techniques for efficient surrogate-assisted optimization, in: *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 2965–2974. doi:[10.1109/SSCI47803.2020.9308465](https://doi.org/10.1109/SSCI47803.2020.9308465).
- 880 [35] L. Van Der Maaten, E. Postma, J. Van den Herik, Dimensionality reduction: a comparative review, *J Mach Learn Res* 10 (2009) 66–71.
- [36] W. H. Lopez Pinaya, S. Vieira, R. Garcia-dias, A. Mechelli, "Autoencoders", Elsevier, 2019, pp. 193–208. doi:[10.1016/B978-0-12-815739-8.00011-0](https://doi.org/10.1016/B978-0-12-815739-8.00011-0).
- 885 [37] L. Agostini, Exploration and prediction of fluid dynamical systems using auto-encoder technology, *Physics of Fluids* 32 (6) (2020) 067103. doi:[10.1063/5.0012906](https://doi.org/10.1063/5.0012906).
- [38] B. Zhao, H. Lu, S. Chen, J. Liu, D. Wu, Convolutional neural networks for time series classification, *Journal of Systems Engineering and Electronics* 28 (1) (2017) 162–169. doi:[10.21629/JSEE.2017.01.18](https://doi.org/10.21629/JSEE.2017.01.18).
- 890 [39] X. Zhao, X. Han, W. Su, Z. Yan, Time series prediction method based on Convolutional Autoencoder and LSTM, in: *2019 Chinese Automation Congress (CAC)*, 2019, pp. 5790–5793. doi:[10.1109/CAC48633.2019.8996842](https://doi.org/10.1109/CAC48633.2019.8996842).
- [40] F. Gers, J. Schmidhuber, F. Cummins, Learning to forget: continual prediction with LSTM, in: *1999 Ninth International Conference on Artificial Neural Networks ICANN 99*. (Conf. Publ. No. 470), Vol. 2, 1999, pp. 850–855 vol.2. doi:[10.1049/cp:19991218](https://doi.org/10.1049/cp:19991218).
- 895 [41] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, W.-c. Woo, Convolutional lstm network: A machine learning approach for precipitation nowcasting, in: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, NIPS'15*, MIT Press, Cambridge, MA, USA, 2015, p. 802810.
- 900 [42] D. Rancourt, S. Ghosh, D. N. Mavris, S. Coggon, A Methodology to Create Approximate Models of Load Envelopes Under Uncertainty. doi:[10.2514/6.2015-3091](https://doi.org/10.2514/6.2015-3091).
- [43] S. Ghosh, D. Rancourt, D. N. Mavris, S. Coggon, Principal Component Analysis Assisted Surrogate Modeling (PCA-SM) of Correlated Loads for Uncertainty Analysis of Design Load Envelopes. doi:[10.2514/6.2015-3092](https://doi.org/10.2514/6.2015-3092).

- 905 [44] I. Tartaruga, J. E. Cooper, P. Sartor, M. H. Lowenberg, Y. Lemmens, Geometrical based method for the uncertainty quantification of correlated aircraft loads, *Journal of Aeroelasticity and Structural Dynamics* 4, No1 (2016).
- [45] I. Tartaruga, J. E. Cooper, P. Sartor, M. H. Lowenberg, Y. Lemmens, S. Coggon, Prediction and uncertainty propagation of correlated time-varying quantities using surrogate models, 910 *CEAS Aeronautical Journal* 7(1) 29–42.
- [46] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* 2 (5) (1989) 359–366. doi:[https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- 915 [47] K. Slawomir, P. D. Anna, Basics of data-driven surrogate modeling, in: *Performance-Driven Surrogate Modeling of High-Frequency Structures*, Springer International Publishing, Cham, 2020, pp. 23–58. doi:[10.1007/978-3-030-38926-0\\_2](https://doi.org/10.1007/978-3-030-38926-0_2).
- [48] E. Torre, S. Marelli, P. Embrechts, B. Sudret, Data-driven polynomial chaos expansion for machine learning regression, *Journal of Computational Physics* 388 (2019) 601–623. doi:<https://doi.org/10.1016/j.jcp.2019.03.039>.
- 920 [49] J. P. Murcia, P. E. Rthor, N. Dimitrov, A. Natarajan, J. D. Srensen, P. Graf, T. Kim, Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates, *Renewable Energy* 119 (2018) 910–922. doi:<https://doi.org/10.1016/j.renene.2017.07.070>.
- [50] S. Chakraborty, S. Adhikari, R. Ganguli, The role of surrogate models in the development of digital twins of dynamic systems, *Applied Mathematical Modelling* 90 (2021) 662–681. 925 doi:<https://doi.org/10.1016/j.apm.2020.09.037>.
- [51] A. Figueroa, M. Göttsche, Gaussian processes for surrogate modeling of discharged fuel nuclide compositions, *Annals of Nuclear Energy* 156 (2021) 108085. doi:<https://doi.org/10.1016/j.anucene.2020.108085>.
- 930 [52] G. F. N. Goncalves, A. Batchvarov, Y. Liu, Y. Liu, L. R. Mason, I. Pan, O. K. Matar, Data-driven surrogate modeling and benchmarking for process equipment, *Data-Centric Engineering* 1 (2020) e7. doi:[10.1017/dce.2020.8](https://doi.org/10.1017/dce.2020.8).
- [53] M. Tahkola, J. Kernen, D. Sedov, M. F. Far, J. Kortelainen, Surrogate modeling of electrical machine torque using artificial neural networks, *IEEE Access* 8 (2020) 220027–220045. doi:[10.1109/ACCESS.2020.3042834](https://doi.org/10.1109/ACCESS.2020.3042834).
- 935 [54] R. Yondo, E. Andrés, E. Valero, A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses, *Progress in Aerospace Sciences* 96 (2018) 23–61. doi:<https://doi.org/10.1016/j.paerosci.2017.11.003>.
- [55] G. Sun, S. Wang, A review of the artificial neural network surrogate modeling in aerodynamic design, *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 233 (16) (2019) 5863–5872. doi:[10.1177/0954410019864485](https://doi.org/10.1177/0954410019864485). 940
- [56] A. Chiplunkar, E. Rachelson, M. Colombo, J. Morlier, Adding flight mechanics to flight loads surrogate model using multi-output gaussian processes, *17th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference* doi:[10.2514/6.2016-4000](https://doi.org/10.2514/6.2016-4000).

- 945 [57] G. Holmes, P. Sartor, S. Reed, P. Southern, K. Worden, E. Cross, Prediction of landing gear loads using machine learning techniques, *Structural Health Monitoring* 15 (5) (2016) 568–582. doi:10.1177/1475921716651809.
- [58] J. Hoole, P. Sartor, J. D. Booker, J. E. Cooper, X. Gogouvis, R. K. Schmidt, Comparison of surrogate modeling methods for finite element analysis of landing gear loads, *AIAA Scitech Forum* (2020). doi:10.2514/6.2020-0681.
- 950 [59] M. Vohra, P. Nath, S. Mahadevan, Y. T. Tina Lee, Fast surrogate modeling using dimensionality reduction in model inputs and field output: Application to additive manufacturing, *Reliability Engineering & System Safety* 201 (2020) 106986. doi:https://doi.org/10.1016/j.res.2020.106986.
- [60] M. Guo, J. S. Hesthaven, Reduced order modeling for nonlinear structural analysis using gaussian process regression, *Computer Methods in Applied Mechanics and Engineering* 341 (2018) 807–826. doi:https://doi.org/10.1016/j.cma.2018.07.017.
- [61] A. Bertram, C. Othmer, R. Zimmermann, Towards real-time vehicle aerodynamic design via multi-fidelity data-driven reduced order modeling, 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference doi:10.2514/6.2018-0916.
- 960 [62] D. Viúdez-Moreiras, M. Martin, R. Abarca, E. Andrés, J. Ponsin, F. Monge, Surrogate modeling for the main landing gear doors of an airbus passenger aircraft, *Aerospace Science and Technology* 68 (2017) 135–148. doi:https://doi.org/10.1016/j.ast.2017.04.021.
- [63] R. Pulch, M. Youssef, Machine learning for trajectories of parametric nonlinear dynamical systems, *Journal of Machine Learning for Modeling and Computing* 1 (2020) 75–95. doi:10.1615/JMachLearnModelComput.2020034093.
- 965 [64] J. Hesthaven, S. Ubbiali, Non-intrusive reduced order modeling of nonlinear problems using neural networks, *Journal of Computational Physics* 363 (02 2018). doi:10.1016/j.jcp.2018.02.037.
- [65] M. Salvador, L. Dede', A. Manzoni, Non intrusive reduced order modeling of parametrized PDEs by kernel POD and neural networks, arXiv:2103.17152 (2021).
- 970 [66] X. He, Q. He, J.-S. Chen, Deep autoencoders for physics-constrained data-driven nonlinear materials modeling, *Computer Methods in Applied Mechanics and Engineering* 385 (2021) 114034. doi:https://doi.org/10.1016/j.cma.2021.114034.
- [67] L. Zamparo, Z. Zhang, Deep autoencoders for dimensionality reduction of high-content screening data, *ArXiv abs/1501.01348* (2015).
- 975 [68] A. Kashefi, D. Rempe, L. J. Guibas, A point-cloud deep learning framework for prediction of fluid flow fields on irregular geometries, *Physics of Fluids* 33 (2) (2021) 027104. doi:10.1063/5.0033376.
- [69] K. Fukami, R. Maulik, N. Ramachandra, K. Fukagata, K. Taira, Global field reconstruction from sparse sensors with voronoi tessellation-assisted deep learning, *Nature Machine Intelligence* 3(11) 945–951.
- 980

- [70] H. Gao, L. Sun, J.-X. Wang, Phygeonet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state pdes on irregular domain, *Journal of Computational Physics* 428 (2021) 110079. doi:<https://doi.org/10.1016/j.jcp.2020.110079>.
- 985 [71] F. Ogoke, K. Meidani, A. Hashemi, A. B. Farimani, Graph convolutional networks applied to unstructured flow field data, *Machine Learning: Science and Technology* 2 (4) (2021) 045020. doi:[10.1088/2632-2153/ac1fc9](https://doi.org/10.1088/2632-2153/ac1fc9).
- [72] R. Pascanu, T. Mikolov, Y. Bengio, On the difficulty of training recurrent neural networks, Vol. 28, No3 of *Proceedings of Machine Learning Research*, PMLR, Atlanta, Georgia, USA, 2013, pp. 1310–1318.
- 990 [73] A. Moghar, M. Hamiche, Stock market prediction using lstm recurrent neural network, *Procedia Computer Science* 170 (2020) 1168–1173, the 11th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops. doi:<https://doi.org/10.1016/j.procs.2020.03.049>.
- 995 [74] B. Lehner, J. Schluter, G. Widmer, Online, loudness-invariant vocal detection in mixed music signals, *IEEE/ACM Trans. Audio, Speech and Lang. Proc.* 26 (8) (2018) 13691380. doi:[10.1109/TASLP.2018.2825108](https://doi.org/10.1109/TASLP.2018.2825108).
- [75] Y. Zeng, J. Zhang, B. Starly, Recurrent neural networks with long term temporal dependencies in machine tool wear diagnosis and prognosis, *SN Applied Sciences* 3 (04 2021). doi:[10.1007/s42452-021-04427-5](https://doi.org/10.1007/s42452-021-04427-5).
- 1000 [76] G. Lan, Q. Li, N. Cheng, Remaining useful life estimation of turbofan engine using lstm neural networks, in: 2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC), 2018, pp. 1–5. doi:[10.1109/GNCC42960.2018.9019107](https://doi.org/10.1109/GNCC42960.2018.9019107).
- 1005 [77] R. Zhang, Z. Chen, C. Su, J. Zheng, O. Bykztrk, H. Sun, Deep long short-term memory networks for nonlinear structural seismic response prediction, *Computers & Structures* 220 (2019) 55–68. doi:[10.1016/j.compstruc.2019.05.006](https://doi.org/10.1016/j.compstruc.2019.05.006).
- [78] K. Lee, J. Kou, W. Zhang, Deep neural network for unsteady aerodynamic and aeroelastic modeling across multiple mach numbers, *Nonlinear Dynamics* 96 (2019) 1–21. doi:[10.1007/s11071-019-04915-9](https://doi.org/10.1007/s11071-019-04915-9).
- 1010 [79] Y. Wang, A new concept using lstm neural networks for dynamic system identification, in: 2017 American Control Conference (ACC), 2017, pp. 5324–5329. doi:[10.23919/ACC.2017.7963782](https://doi.org/10.23919/ACC.2017.7963782).
- [80] P. R. Vlachas, W. Byeon, Z. Y. Wan, T. P. Sapsis, P. Koumoutsakos, Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 474 (2213) (2018) 20170844. doi:[10.1098/rspa.2017.0844](https://doi.org/10.1098/rspa.2017.0844).
- 1015 [81] G. E. Hinton, R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* 313 (5786) (2006) 504–507. doi:[10.1126/science.1127647](https://doi.org/10.1126/science.1127647).

- 1020 [82] Y. Huang, C. H. Chen, C. J. Huang, Motor fault detection and feature extraction using rnn-based variational autoencoder, *IEEE Access* 7 (2019) 139086–139096. doi:10.1109/ACCESS.2019.2940769.
- [83] J. Yang, S. Zhang, Y. Xiang, J. Liu, J. Liu, X. Han, F. Teng, Lstm auto-encoder based representative scenario generation method for hybrid hydro-pv power system, *IET Generation, Transmission & Distribution* 14 (2020) 5935–5943(8). doi:10.1049/iet-gtd.2020.0757.
- 1025 [84] A. Sagheer, M. Kotb, Unsupervised pre-training of a deep lstm-based stacked autoencoder for multivariate time series forecasting problems, *Scientific Reports* 9 (2019) 19038. doi:10.1038/s41598-019-55320-6.
- [85] H. D. Nguyen, K. P. Tran, S. Thomassey, M. Hamad, Forecasting and anomaly detection approaches using lstm and lstm autoencoder techniques with the applications in supply chain management, *International Journal of Information Management* (11 2020). doi:10.1016/j.ijinfomgt.2020.102282.
- 1030 [86] T. Wong, Z. Luo, Recurrent auto-encoder model for large-scale industrial sensor signal analysis, in: P. Elias, J. Chrisina (Eds.), *Engineering Applications of Neural Networks*, Springer International Publishing, Cham, 2018, pp. 203–216. doi:https://doi.org/10.1007/978-3-319-98204-5\_17.
- 1035 [87] Q. Meng, D. R. Catchpoole, D. Skillicom, P. J. Kennedy, Relational autoencoder for feature extraction, 2017 International Joint Conference on Neural Networks (IJCNN) (2017) 364–371doi:10.1109/IJCNN.2017.7965877.
- [88] R. G. Hefron, B. J. Borghetti, J. C. Christensen, C. M. S. Kabban, Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation, *Pattern Recognition Letters* 94 (2017) 96–104. doi:https://doi.org/10.1016/j.patrec.2017.05.020.
- 1040 [89] M. Abadi, P. Barham, J. Chen, Z. e. a. Chen, Tensorflow: A system for large-scale machine learning, in: *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, OSDI'16*, USENIX Association, USA, 2016, p. 265283.
- 1045 [90] M. D. McKay, Latin hypercube sampling as a tool in uncertainty analysis of computer models, in: *Proceedings of the 24th Conference on Winter Simulation, WSC '92*, Association for Computing Machinery, New York, NY, USA, 1992, p. 557564. doi:10.1145/167293.167637.
- 1050 [91] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, et al., Scipy 1.0: fundamental algorithms for scientific computing in python, *Nature Methods* 17 (2020) 261–272. doi:10.1038/s41592-019-0686-2.
- [92] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings (2015).
- 1055 [93] M. Milano, P. Koumoutsakos, Neural network modeling for near wall turbulent flow, *Journal of Computational Physics* 182 (1) (2002) 1–26. doi:https://doi.org/10.1006/jcph.2002.7146.

- 1060 [94] T. Murata, K. Fukami, K. Fukagata, Nonlinear mode decomposition with convolutional neural networks for fluid dynamics, *Journal of Fluid Mechanics* 882 (2020) A13. doi:10.1017/jfm.2019.822.
- [95] K. Fukami, K. Hasegawa, T. Nakamura, M. Morimoto, K. Fukagata, Model order reduction with neural networks: Application to laminar and turbulent flows, *SN Computer Science* 2 (1) (2021) 1–16. doi:10.1007/s42979-021-00867-3.
- 1065 [96] P. Stojakovi, K. Velimirovi, B. Rasuo, Power optimization of a single propeller airplane take-off run on the basis of lateral maneuver limitations, *Aerospace Science and Technology* 72 (2018) 553–563. doi:https://doi.org/10.1016/j.ast.2017.10.015.
- [97] P. Stojakovic, B. Rasuo, Single propeller airplane minimal flight speed based upon the lateral maneuver condition, *Aerospace Science and Technology* 49 (2016) 239–249. doi:https://doi.org/10.1016/j.ast.2015.12.012.
- 1070 [98] P. Stojakovic, B. Rasuo, Minimal safe speed of the asymmetrically loaded combat airplane, *Aircraft Engineering and Aerospace Technology: An International Journal* 88 (2016) 42–52. doi:https://doi.org/10.1016/j.ast.2015.12.012.
- [99] M. Bajovic, M. Zivanovic, B. Rasuo, P. Stojakovic, New general approach to airplane rotation analysis, *Japan Society of Aeronautical Space Sciences Transactions* 53 (2010) 130–137. doi:10.2322/tjsass.53.130.
- 1075 [100] L. Prechelt, Automatic early stopping using cross validation: quantifying the criteria, *Neural Networks* 11 (4) (1998) 761–767. doi:https://doi.org/10.1016/S0893-6080(98)00010-0.
- [101] J. Wright, J. Cooper, *Introduction to Aircraft Aeroelasticity and Loads: Second Edition*, Wiley, 2015, Ch. 16, pp. 367–394. doi:10.1002/9781118700440.
- 1080 [102] P. Wang, N. C. Bouaynaya, L. Mihaylova, J. Wang, Q. Zhang, R. He, Bayesian neural networks uncertainty quantification with cubature rules, in: *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–7. doi:10.1109/IJCNN48605.2020.9207214.
- [103] D. P. Kingma, M. Welling, Auto-encoding variational bayes, in: Y. Bengio, Y. LeCun (Eds.), *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.
- 1085 [104] G. Karniadakis, Y. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, *Nature Reviews Physics* 3 (2021) 422440. doi:10.1038/s42254-021-00314-5.
- [105] S. Qian, G. Li, W.-M. Cao, C. Liu, S. Wu, H. S. Wong, Improving representation learning in autoencoders via multidimensional interpolation and dual regularizations, in: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization*, 2019, pp. 3268–3274. doi:10.24963/ijcai.2019/453.
- 1090