DINAME 2023 - Proceedings of the XIX International Symposium on Dynamic Problems of Mechanics,
A.S. De Paula, A.T. Fabro and M.R. Machado (Editors), ABCM.
Pirenópolis, Go, Brazil, February 26th to March 3rd, 2023

Digital Twin of an Offshore Riser Systems for Time Series Prediction Using Deep Learning Models

Stanley Washington Ferreira de Rezende[1], Leandro Soares Silva[1], Aldemir Ap. Cavalini Jr.[1], Jose dos Reis Vieira de Moura Jr.[2] and Valder Steffen Jr.[1]

[1] LMEst – Structural Mechanics Laboratory, Federal University of Uberlandia, School of Mechanical Engineering, Av. João Naves de Ávila, 2121, Uberlândia, MG, 38408-100, Brazil.

[2] Federal University of Catalao, Mathematics and Technology Institute, Av. Dr. Lamartine Pinto de Avelar, 1120, Vila Chaud, Catalão, GO, 75704-020, Brazil

Abstract: Due to the shortcomings of numerical models and their expensive processing requirements, the use of deep learning models for time series prediction has become feasible. For this reason, the implementation of an API (Application Programming Interface) of the Digital Twin to predict time-domain dynamic responses of a riser structure is proposed in the present contribution. The Long-Short Term Memory (LSTM) and Multilayer Perceptron (MLP) neural networks, employed as prospective prediction models, constitute the foundation of the Digital Twin. The oil platform's actual displacements and rotations caused by the sea's waves are the time series used to create the dynamic reactions of the riser. The forces produced by a finite element model (FEM), using the ANFLEX software developed by CENPES/PETROBRAS, were used as the training output time series. The output of the two neural networks was evaluated and compared with each other both for the training series and for the prediction of series not yet seen by the models. Four metrics were calculated to facilitate comparison, and signal comparisons versus time graphs were plotted.

Keywords: Digital Twin, Oil and Gas Industry, Deep Learning Models, Offshore Riser Systems

INTRODUCTION

The development of offshore oil exploration activities in deep and ultra-deep waters has continuously increased in recent years (Costa *et al.*, 2003). This task usually involves the use of floating production systems (FPS), kept in position by mooring lines and connected to subsea equipment (such as manifolds and wellheads) by production risers, which transport oil production from the drilling wells to the platform or vessel (De Pina *et al.*, 2011; MA *et al.*, 2013).

However, the harsh circumstances imposed on risers as water depth increases can result in cyclic efforts that can damage the component over time. In this sense, for safety and economic reasons, it is essential to monitor the structural behavior of production risers (Clarke *et al.*, 2011).

Monitoring a mechanical system's structural integrity aims to forecast the physical behavior of an actual structure, which depends on several factors; each interaction between its constituent parts and particles, for example, directly influences how the structure is currently functioning (Farrar and Worden, 2007). Therefore, the modeling to be considered of the structure under study must be done by replicating the characteristics and parameters of the components involved in a virtual environment, in which different usability scenarios are thoroughly investigated to predict the future dynamic (or static) behavior of the system (Kritzinger *et al.*, 2018).

The term "digital twin" refers to virtualizing a physical system's behavior in a digital model. And in that context, various types of probabilistic simulations and experimental data are used to ensure that the predictions and insights made using the digital model are as accurate as possible (Soderberg *et al*., 2017; Rangel *et al.*, 2021; Priyanka *et al.*, 2022).

Thus, using several simulations and tests on the digital twin model, a viable decision can be made based on available resources and systems. For that, still in the design step, models are used based on specific information about the system's operation and simulated loads for the design circumstances (standardized according to international standards) to predict the useful life of the production risers. Then, real measurement data are continuously used to reduce the uncertainties of the developed models, following the entire useful life of the risers and gradually improving the analyzes performed on them (Droder *et al.*, 2018; Priyanka *et al.*, 2022).

Many solutions have been used to model flexible risers, including finite element models (FEM) and analytical methods (Santillan and Virgin, 2011). However, the complexity of numerical models makes their application unfeasible due to the high computational cost and the fact that the problem is mathematically ill-posed (Rangel *et al.*, 2021).

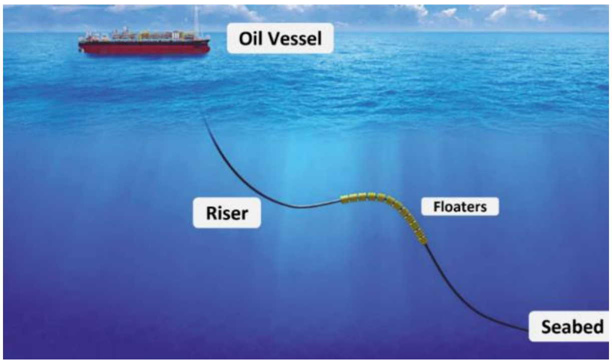
At the same time, deep learning (DL) models currently represent one of the main ways of describing a physical problem in a virtual system, and this fact is due to the ability of such models to integrate with different cutting-edge technologies, fast processing and ease of use (Harrison, 2019; Priyanka *et al.*, 2022).

Therefore, DL algorithms can be considered useful for the formulation of digital twin models, and, in this sense, the objective of this article is to present the functioning of an API software intended to predict the temporal responses of a flexible riser, which result from to the displacements and rotations imposed on the oil platform on the sea surface.

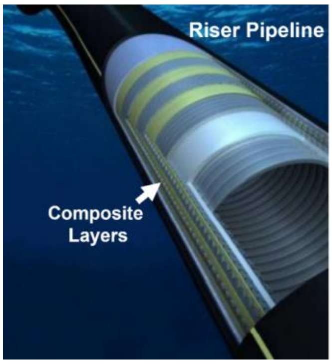
In this way, two distinct artificial neural networks (ANNs), namely the MLP (Multilayer Perceptron) and the LSTM (Long-Short Term Memory), were used as a prediction model, and other machine learning subprocesses reinforced the whole procedure; such as the clusterization of the series by the k-means method, the optimization of the neural architectures by the differential-evolution optimization method (DE) and others.

OFFSHORE RISER SYSTEMS MODELING TYPES

Control of the dynamics of a marine riser, a tubular structure that connects an offshore oil and gas platform to a well on the seabed (Fig. 1a), has become essential for ocean and control engineers due to the need for the production of oil and gas from the seabed (Do and Pan, 2009). As a result, production risers must, therefore, be built with high-resistance materials to offer a constant flow of production and flow, in addition to being resistant to non-linear stresses brought about by the environment in which they are inserted, such as waves and ocean currents, winds and others (Clarke *et al.*, 2011). As a result, offshore production risers are an example of composite structures (Fig. 1b), formed by several independent layers of material, which are loosely connected to interact with each other (Costa *et al.*, 2003).



a) Offshore production system components.



b) Flexible riser composition.

Figure 1 – Description of the offshore production system.

At the same time, flexible risers are considered critical structures due to their activity and connection with environmental and financial risks. Consequently, fatigue life studies of this type of structure are constantly needed, and its modeling so far is commonly based on finite element models (Costa *et al.*, 2003; Riveros *et al.*, 2007; Clarke *et al.*, 2011).

Fluid Dynamics and Immersed Structures are branches of Mechanical Engineering with which the behavior of fluid-structural interactions is studied, aiming to improve several processes in the industry and understand the phenomena in nature. Therefore, regarding determining the useful life of production risers, this phenomenon can be studied in two ways: material experimentation and virtual experimentation. On both fronts, it is necessary to physically model the problem under analysis and set up experimental platforms that represent the physics of the problem to be studied. (Kim *et al.*, 2021).

Therefore, the physical modeling of a riser consists of evaluating the problem of interest and determining the physical premises that interfere with its analysis, such as boundary conditions, imposed forces, connectivity between elements, and others. Some of these properties, in turn, can be difficult to measure and model. In contrast, the mathematical modeling of a mechanical system consists of obtaining the differential, integral, and differential-integral equations that model the associated physics and then using appropriate numerical methods to discretize such equations (Riveros *et al.*, 2007; Kim *et al.*, 2021).

Thus, the motivation for developing this article is that the computational cost of simulations commonly related to the design of immersed structures is high. This problem becomes even more significant if it is necessary to apply processes of optimization and readjustment of the project, such as the definition of the number of finite elements used in FEM studies.

For these and other reasons, in the present work, a computational tool was developed in python language dedicated to the automatic construction of ANNs models that represent the dynamic behavior of flexible risers with a specific efficiency level.

DIGITAL TWIN API FORMULATION

In general, digital twin models must be flexible enough to adapt to environmental changes and virtually capture the behavior of a physical system under study. Thus, for the formulation of this digital twin API, two aspects of its use were individually implemented, the first aimed at creating new digital models and the second related to the application and adaptation of pre-existing models. Then, the neural models predict efforts in each riser element.

Fig. 2, in this sense, presents the processing flow considered during the first scenario of using the digital twin API implemented here.

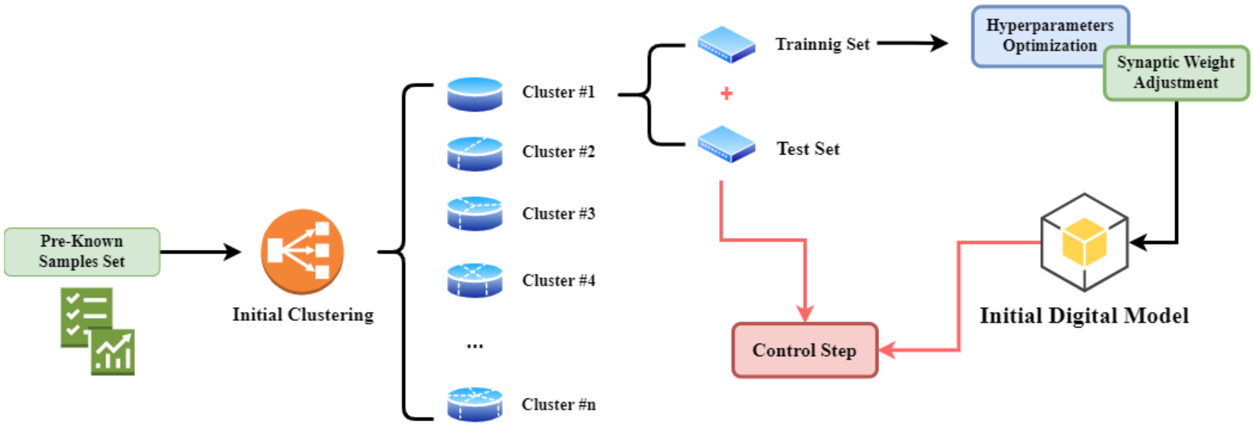


Figure 2 – Diagram of the first Digital Twin API usage step.

It is important to note that both neural models currently implemented in the API employ the supervised training method. In this way, the first API usage scenario starts with a set of previously known samples (that is, the target values of each sample are already known in advance). Then an initial grouping of these samples into clusters is performed. This clustering procedure, in turn, is committed to reducing the visible variability of the data by the digital models.

Subsequently, two subsets of samples are then defined from the dataset of each previously delimited cluster, one of which is used for creating and training the current digital model and the other used for controlling the digital twin API.

On the other hand, the training process of neural models occurs in two steps, the first aiming to optimize the neural architecture to be considered a digital model and the second based on adjusting its synaptic weights and bias. Therefore, in the first process, the number of neural layers, the number of neurons in each layer, the dropout rates, and the weights optimization model are used as design variables of the differential evolution optimization method.

The objective function used in the DE method is given as the minimization of the prediction error on a validation dataset of 30% of the training samples set. In contrast, synaptic weights are adjusted using gradient descent optimization methods, such as the Adam method, RMSprop, SGD, etc.

After building and training the neural model, the initial API control and validation step is carried out. For this, the newly created digital model is then evaluated based on the training dataset's metric R2 (Eq. 1). If the R^2 metric between the actual value (t_i) of a time series and the predicted value (y_i) by the initial digital model is above a pre-established threshold value, the considered series is then kept in the group. Otherwise, it is transferred for a transient set.

$$R^2 = 1 - \frac{\sum_{i=1}^{n\_points}(t_i - y_i)^2}{\sum_{i=1}^{n\_points}(t_i - \bar{t}_i)^2} \tag{1}$$

Therefore, as the time series are removed from the cluster, the ANN retraining is performed based on the maintained series, and a new digital model is created. It should also be noted that with the update of the series belonging to the cluster, the group's centroid is also updated, as well as the other components measured by the entire API.

Finally, the resulting digital model is also evaluated in the test dataset. If any sample from this set does not meet the pre-established threshold value, it is also taken to the transient dataset. However, there is no model retraining at this step since such samples will only serve as a control for the second phase of the API.

Furthermore, having completed the entire control procedure of the training and test samples, three novelty detection models (Local Outlier Factor - LOF, Isolation Forest – IF, and Copula-Based Outlier Detection - COPOD) are applied to the maintained samples. This procedure is performed to create the novelty detection models used in the second API usage scenario (Fig. 3).

After the API creates the neural models, these can be used to support fatigue life analysis. In this way, as new experimental samples of the behavior of the offshore platform become available, they can also be analyzed by the previously created digital models. Several characteristics related to the useful life of the riser in question can also be deduced at once. The second part of the digital twin API implementation is used, as shown in Fig. 3.

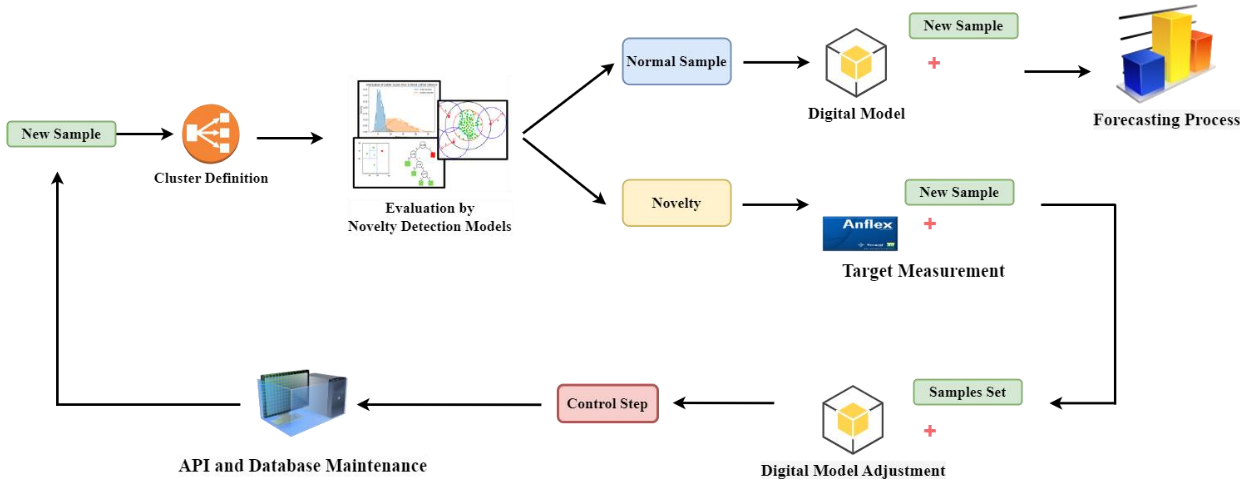


Figure 3 – Diagram of the second Digital Twin API usage step.

As can be seen, when a new simulation of displacements and rotations of the offshore platform is presented to the API, it is then grouped according to the pre-existing clusters in the system's database. Then, all three novelty detection models belonging to the defined collection are applied to the input time series of the model, categorizing the evaluated simulation as a new sample or not.

Suppose the current sample is statistically considered standard selection (i.e., it belongs to the evaluated dataset). In that case, the stored digital model is loaded, and the procedure for predicting the efforts imposed in the riser for the element under analysis is performed. Therefore, this prediction procedure is characterized as the ideal flow of the API. The implementation makes it possible to predict the time series of efforts in the riser with the lowest possible computational cost.

On the other hand, if the model input series are considered a new sample, they are transferred to a new simulation in the ANFLEX software to define their respective target values. Then, the neural architecture is optimized, and the neural weights and centroids of the clusters are also adjusted as part of a new step of updating the digital model. If there is any change in the digital model, the same control procedure as in the previous step is performed again, and consequently, the entire database is updated.

In this way, when using the two implementations of this digital twin API, it is possible to predict the vibratory behavior of flexible risers in real offshore installations, helping fatigue life studies, reducing costs, and avoiding setbacks.

TIME SERIES PREDICTION – CASE OF STUDY

To verify the efficiency of the developed methodology, in the present work, a performance study of the digital twin API was carried out to predict the temporal response data of a real riser structure (already under implementation). In this sense, a set of 1100 samples of displacements and rotations of the offshore platform was considered to create and use the MLP and LSTM digital models.

Each sample considered, in turn, was formed by seven-time series of 18,000 points ($\approx 3{,}500s$) representing longitudinal displacements (X, Y, Z), rotations (R_X, R_Y, R_Z), and efforts (F_X) estimated in the region of connection between the riser and the offshore platform, that is, in the TOP part of the system (Fig. 4).

It is also worth mentioning that while the input values (displacements and rotations in the X, Y, and Z axes) of the model were obtained experimentally, the output values (F_X) considered as reference values in training were obtained through different simulations of the system in the ANFLEX software. On the other hand, such simulations were limited to analyzing a slow wave configuration, placing floats on a long section of the riser, and exhibiting part of the tube's buoyancy.

The presence of a floating region (called SAG/HOG region) in the riser attenuates the loads induced by waves transferred from the FPSO (Floating Production Storage and Offloading) to the part in contact of the riser with the seabed (TDP). In this way, the mechanical system used in the present contribution can be interpreted as the union of 3 catenaries. The central section is an inverted catenary, as shown in Fig. 4.

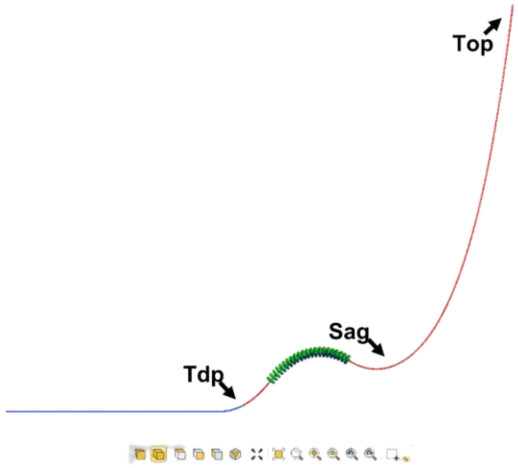


Figure 4 – Physical model diagram obtained by ANFLEX software.

Subsequently, from the total of samples evaluated, 100 randomly selected samples were used to apply the first scenario of using the API, which is the construction of digital models. Therefore, the k-means clustering method was then used to divide the 100 samples into ten initial clusters, and, for each collection considered, two subsets of data were again chosen at random, representing 80% of the samples destined for training and 20% for testing and model control steps. Fig. 5 shows the distribution of samples in the 10 clusters delimited by the k-means method.

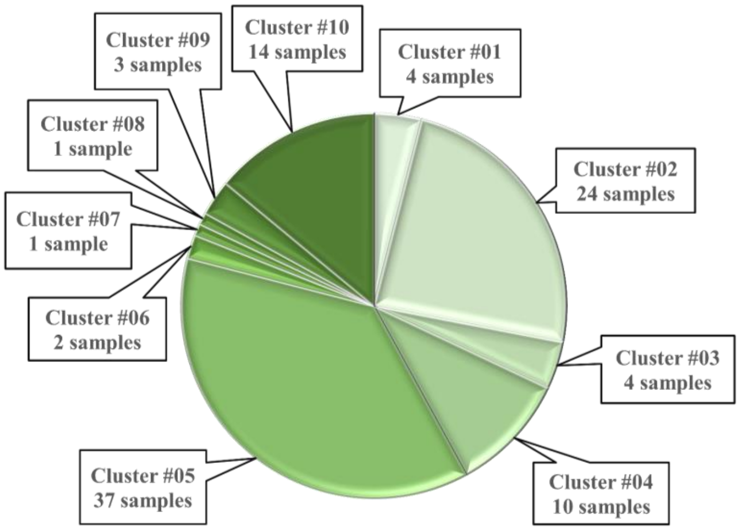


Figure 5 – Samples distribution in the clusters used to create the digital models.

It can be seen in Fig. 5 that the number of samples in each group did not follow a homogeneous pattern, which was expected since the set of samples considered was chosen randomly. However, even if some clusters have smaller initial samples, they can still be regarded as representative digital models of the riser structure. Consequently, they were still considered during the application of the second scenario of API usage.

On the other hand, it is worth mentioning that together with the training, the neural architectures were optimized to find the set of hyperparameters that would obtain the best training for each model/cluster. In this sense, the optimized hyperparameters were the number of hidden layers (n_l), the number of neurons in the first hidden layer (n_e), the dropout rate (d_r), and the optimizer method. In turn, the domains of these search variables were given according to Tab. 1 for the present case study.

Table 1 – Domain of the search variables considered during the ANNs hyperparameter optimization process.

Hyperparameter	Search Domain
n_l	$1 \leq n_l \leq 6$
n_e	$422 \leq n_e \leq 1500$
d_r	$0.1 \leq d_r \leq 0.5$
Optimizer Method	Adam, Nadam, Adamax

Furthermore, a total of 70 epochs was considered during training to modify the synaptic weights in each iteration of the DE optimization. The MSE loss function was used to adjust the weights and biases of the ANNs. Then, the API control flow was applied, and the R^2 metric values associated with each series were measured and compared with a threshold value equal to 0,8. Therefore, the series whose metric values were lower than this was eliminated from the cluster, and the developed digital model was updated.

Finally, the three novelty detection models mentioned above were built using the maintained training series set, and the remaining samples (1000 simulations) were subsequently evaluated using the second scenario of API usage, which is the direct prediction of the efforts imposed in the riser.

Fig. 6, in this sense, presents the prediction results of some series from the clusters 2, 5, and 10 for both ANNs used (MLP on the left and LSTM on the right), with these predictions being acquired after the application of both usage scenarios of the digital twin API implemented here.

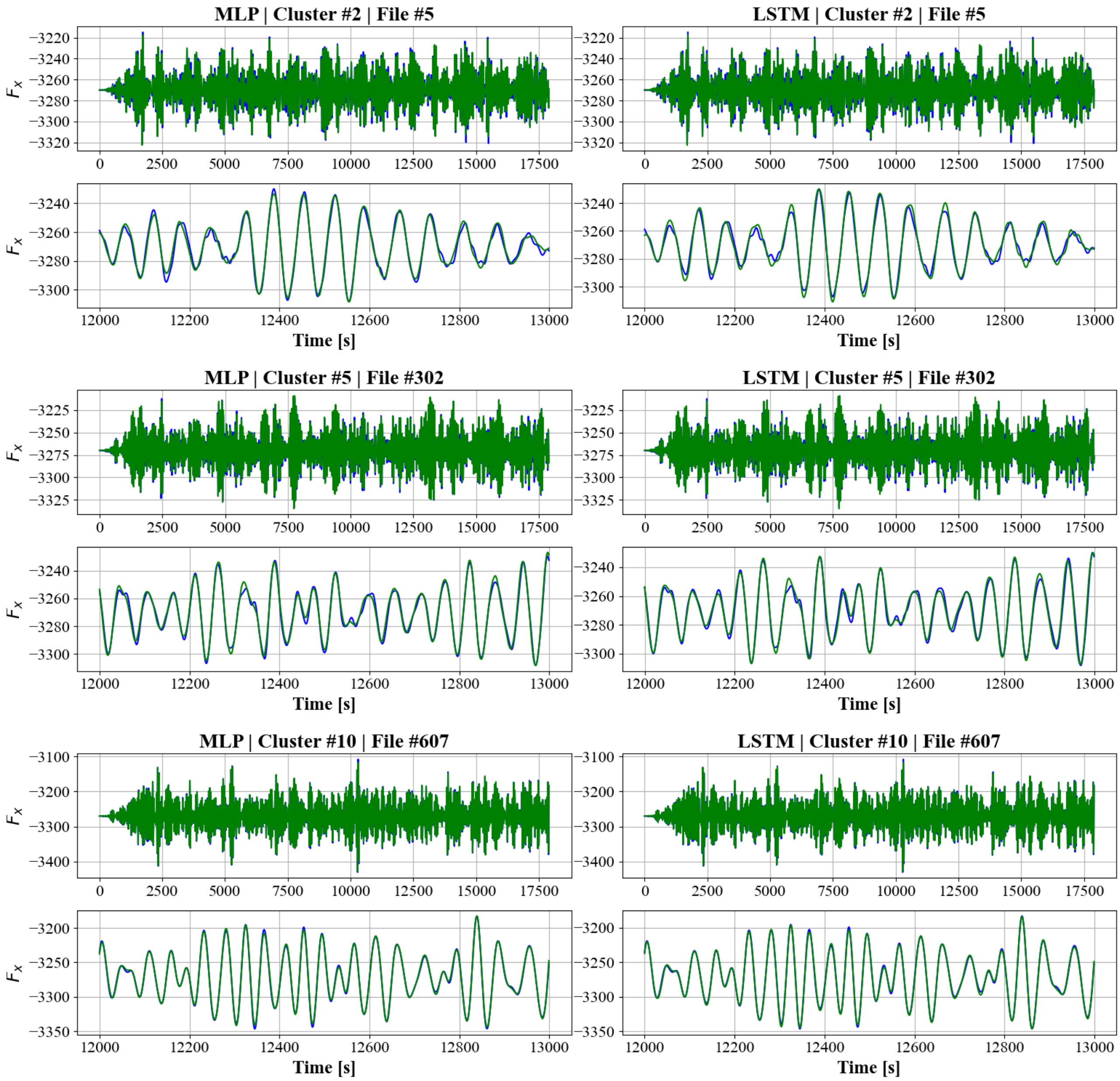


Figure 6 – Result of time series forecasts obtained by each digital model.

In Fig. 6, the blue signals represent the target values (i.e., the effort values measured by the ANFLEX software), while the green signals represent the predicted values by the MLP and LSTM models. Thus, it is noticed that the methodology developed in this work presents good accuracy regarding the estimation of the F_x efforts signals imposed on the TOP region of the riser structure under study.

However, four precision metrics (MAE, MSE, Max Error, and R^2) were calculated between each predicted and expected signal to obtain a quantitative value of this estimate. Therefore, Tab. 2 presents the average values of the metrics achieved for Clusters 2, 5, and 10 for all evaluated series.

Table 2 – Mean metric values obtained for each model's Clusters 2, 5, and 10.

Cluster	Digital Model	MAE	MSE	Max Error	R^2
2	MLP	0.999509	1.636112	5.398839	0.998359
	LSTM	1.608648	4.268758	8.133238	0.997285
5	MLP	1.899576	5.799690	9.132399	0.983261
	LSTM	1.426511	3.311107	6.583046	0.996680
10	MLP	2.110137	7.285553	11.240547	0.971145
	LSTM	5.424066	53.531087	59.826616	0.967059

It is worth mentioning that although the values of the Max Error and MSE metrics were high for some data samples, the predictions made by the API can still be considered good, as the values of the R^2 metric approached 1 (ideal condition) for both models, in all clusters considered. In this way, the methodology developed here could be validated, and consequently, this digital twin API becomes useful in fatigue life studies of flexible risers.

CONCLUSION

In this work, the operation of a digital twin API for the prediction of dynamic temporal responses of flexible risers was demonstrated. The API was developed in python language, and two deep learning models (MLP and LSTM) were considered digital predictor models. Other machine learning sub-processes and methods, such as clustering, optimization, and novelty detection, were also incorporated into the API processing flow to improve the sensitivity of the developed methodology.

A flexible riser structure was then instrumented as a case study so that the displacements and rotations of the connected offshore platform could serve as a set of inputs for the digital models. On the other hand, the efforts imposed on the connection zone (TOP) between the riser and the platform were the target series of the models. In this sense, 100 samples were previously simulated in the ANFLEX software, and these data were used as a reference for the training process.

Ten initial clusters were considered during the creation of the neural models. Then, 1000 other samples were evaluated through the second API usage scenario. The results of this step were then compared using temporal response graphs and four precision metrics values. These, in turn, reached an average R^2 metric value greater than 0.97 for all clusters monitored by the API. Therefore, the methodology developed in this contribution becomes useful for studies of the fatigue life of flexible risers.

ACKNOWLEDGMENTS

The authors would like to thank Petrobras, CNPq, FAPEMIG and CAPES (INCT-EIE) for the financial support of the present contribution. The authors also would like to thank Foz do Chapecó, Baesa, Enercan and Ceran for technical and financial support, through the Research and Development project *PD* – 02949 – 3007/2022 – "Solução integrada para o diagnóstico de defeitos, análise dinâmica e monitoramento contínuo de unidades de geradoras francis" with resources from ANEEL's R&D program.

REFERENCES

CLARKE, T., JACQUES, R., BISOGNIN, A., CAMERINI, C., DAMASCENO, S. and STROHAECKER, T. (2011). "Monitoring the structural integrity of a flexible riser during a full-scale fatigue test". *Engineering Structures,* 33(4), 1181-1186.

COSTA, C. H. O., ROITMAN, N., MAGLUTA, C. and ELLWANGER, G. B., (2003). "Crush Behavior of Flexible Risers: Prediction by an Analytical and Experimental Approach". *17th International Congress of Mechanical Engineering* – COBEM.

DE PINA, A.A., ALBRECHT, C.H., LIMA, B.S.L.P., JACOB, B.P. (2011). "Tailoring the particle swarm optimization algorithm for the design of offshore oil production risers", *Optimization and engineering* 12(1), pp. 215–235.

DO, K.D. and PAN, J., (2009). “Boundary control of three-dimensional inextensible marine risers”, *Journal of Sound and Vibration* 327(3-5), pp. 299–321.

DRODER, K., BOBKA, P., GERMANN, T., GABRIEL, F. and DIETRICH, F. (2018). “A machine learning-enhanced digital twin approach for human-robot-collaboration”, *Procedia Cirp* 76, pp. 187–192.

FARRAR, C.R. and WORDEN, K. (2007). “An introduction to structural health monitoring”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365(1851), pp. 303–315.

HARRISON, M. (2019). “Machine Learning–Guia de Referência Rápida: Trabalhando com dados estruturados em Python”, *Novatec Editora*.

KIM, S. W., SÆVIK, S., WU, J. and LEIRA, B. J. (2021). Prediction of deepwater riser VIV with an improved time domain model including non-linear structural behavior. *Ocean Engineering*, 236, 109508.

KRITZINGER, W., KARNER, M., TRAAR, G., HENJES, J. and SIHN, W. (2018), “Digital twin in manufacturing: A categorical literature review and classification”, IFAC- *PapersOnLine* 51(11), pp. 1016–1022.

MA, K. T., SHU, H., SMEDLEY, P., L'HOSTIS, D. and DUGGAL, A. (2013). A historical review on integrity issues of permanent mooring systems. *In Offshore technology conference. OnePetro*.

PRIYANKA, E. B., THANGAVEL, S., GAO, X. Z. and SIVAKUMAR, N. S. (2022). Digital twin for oil pipeline risk estimation using prognostic and machine learning techniques. *Journal of industrial information Integration*, 26, 100272.

RANGEL, G. M., AMARAL, G. A., SOUZA NEVES, K. L., MARTINS, G. R., NETO, A. G., JÚNIOR, L. A. G. B., FRANZINI, G. R., DOTTA, R. and MALTA, E. B. (2021). Methodology for Obtaining a Digital Twin for a FPSO Mooring System. *26th International Congress of Mechanical Engineering* – COBEM.

RIVEROS, C., UTSUNOMIYA, T., MAEDA, K. and ITOH, K. (2007). Vibration-based damage detection in flexible risers using time series analysis. *Doboku Gakkai Ronbunshuu A*, 63(3), 423-433.

SANTILLAN, S. AND VIRGIN, L. (2011). “Numerical and experimental analysis of the static behavior of highly deformed risers”, *Ocean Engineering* 38(13), pp. 1397–1402.

SODERBERG, R., WARMEFJORD, K., CARLSON, J.S. and LINDKVIST, L. (2017). “Toward a digital twin for real-time geometry assurance in individualized production”, *CIRP annals* 66(1), pp. 137–140.