Neural network modeling of the nonlinear dynamic structural offshore system with hysteresis

D. M. Rocha[1], N. F. F. Ebecken[2], L. P. Calôba[2] & D. L. Kaiser[1]
[1]*PETROBRAS S.A., Brazil*
[2]*COPPE/Federal University of Rio de Janeiro, Brazil*

Abstract

This paper proposes an empirical modeling of the system formed by the riser-platform connection, in deep water. This connection has the objective of minimizing the acting bending moment, possesses high complexity and high-criticity due to economic and environmental consequences from its fault. The main element in the joint is made of elastomeric material, which reveals nonlinear hysteresis. In addition, this whole connection system presents nonlinearities due to the action of dynamic loading and large motions. TDNN and Recurrent Neural Networks (RNN) are being investigated since they possess the ability to model nonlinear hysteretic behaviors and also dynamic systems. Simulation results have confirmed that RNN is the one that presents the best representation of the system studied. Emphasis shall be given to the additional difficulties, which arise from the utilization of real data in the modeling process for this system.
Keywords: Recurrent Neural Network, NARX, hysteresis, Flexjoint SCR.

1   Introduction

The specific system under study is that which makes the connection between the riser and the platform, through a connector called *Flexjoint*. The Flexjoint (see Figure 1) is installed on the SCR (Steel Catenary Riser) top, and through an elastomeric joint it allows the SCR to bend when submitted to dynamic environmental loading of wave, wind and current. The elastomer plays an important and complex role and any deterioration of the elastomer can affect the

performance of the equipment, although it is not possible to conclude whether such deterioration occurs in regards to an increase or reduction in stiffness. However, it is predictable that, if the connection between the riser and the platform becomes rigid, the fatigue damage would become more severe close to the top of the riser , thereby reducing the fatigue life.
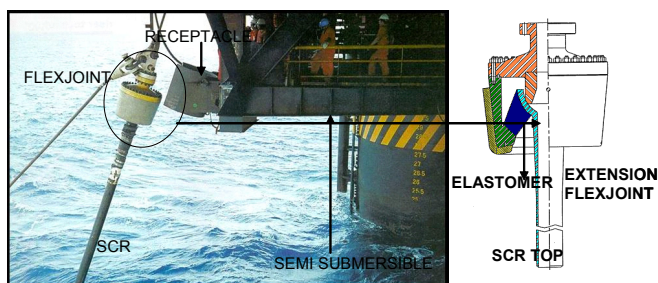

Figure 1:   Flexjoint – platform PXVIII – Brazil Campos Basin.

The purpose of this study is to obtain a more thorough understanding of the hysteretic behavior of the referred joint.

The mathematical formulation of a model based on physical principles and laws that represent the proposed system is very complex, due to the non-linear characteristic of the constituting laws, but especially due to the fact that the elastomer exhibits hysteresis.

The available data and the development of Systems Identification techniques led this work to an empiric modeling of the dynamic behavior of the elastomer. Artificial Neural Networks (ANN) was chosen as the basis of this study.

2   Dynamic neural network theory

Traditionally, the ANN are widely applied to static problems, such as the classification and pattern recognition. Such networks are called static ANN, and they simulate long term memory properties of the human being memory system. Some examples are the MLP (Multilayer Perceptron) and RBF (Radial Basin Function) networks. However, these networks are not capable of representing temporal, dynamic situations, or that may present an order in the set of data.

The hysteresis is a dynamic phenomenon, and therefore its main characteristic depends on the past. Fundamentally, a neural network needs some form of memory to model this type of phenomenon. This memory is known as a short term memory, and it is introduced in the neural network by means of time delays and/or through feedback [1]. These networks can be divided in two groups: the networks with external and with internal dynamics.

The former uses input time delays and/or feedback from the output layer to the input layer. Some examples are the TDNN (Time Delay Neural Network) network: the NOE (Nonlinear output error) network and the NARX (Nonlinear

Autoregressive with exogenous input) networks, also known as parallel (PM) and series-parallel model (SPM), respectively [2].

In the latter, the past values of the input and the output are not reintroduced into the network input, which means that no knowledge a priori is used. The dynamics are inserted by the so-called context memories, and occur through the feedback of the hidden layers. A classical example of this type of architecture is the Elman network [3].

Several studies in many areas propose different architectures for the hysteresis empiric modeling by using ANN. In [4], the proposal is a hybrid solution, called by the author as partially recurrent TDNN, for modeling the stress-strain behavior of geomaterials submitted to cyclical loading. In function of such loading, the geomaterials show hysteresis. The investigation performed by [5] confirms that the sequential RNs with feedback are more effective than the conventional RNs with no feedback to simulate the stress-strain relation of the soil. In problem solving related to water (stage-discharge) the effectiveness of an ANN with a radial basis function and MLP were explored in [6] and [7] respectively. In this paper, the networks MLP, TDNN, NOE and NARX will be analyzed.

3   Data

The data used to generate the Flexjoint behavior model originate from two sources: the monitoring process of the Flexjoint itself and laboratory tests of the elastomer. The following sections describe these two set of data.

3.1 Monitored data

The monitored data are random signals over time, discrete and originating from a monitoring process of a semi-submersible oil/gas platform, PXVIII, located in the Campos Basin, in Rio de Janeiro – Brazil [8]. The relevant variables to this study are as follows: Bending moment (NUM0 and NUM90), towards the riser's catenary and perpendicular (0°, 90°); Axial effort (NUT); Relative angle between the riser and the platform as defined in the riser's catenary plain (APR0), and relative angle between the riser and the platform as defined in the perpendicular plain to the catenary (APR90).

The strain-gage was the sensor adopted for measuring forces and the inclinations have been obtained by means of an inclinometers.

The period that has been taken into account was from July 18, 2001 to November 20, 2001, as it is an acquisition period with a smaller occurrence of acquisition errors, noise and lack of data. The file is composed of 948 discrete series, each one representing a sea state. In short, at least three basic information are necessary in order to specify the conditions of a certain sea state: wave height, period and direction. During the series period, and consequently the sea state, the sea is described as a random, ergodic and stationary process.

A 30 minute period was defined as the time to satisfy the above conditions (ergodicity and stationarity). So, for each sequence, records were made

approximately every 30 minutes every 3 hours, 24 hours a day (Figure 2). The sampling frequency was 1 Hz (1s), totalizing around 1800 points per sequence [8]. This period was determined by the capacity of the data acquisition system.
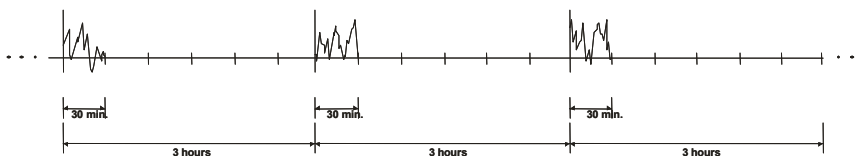

Figure 2:   Illustrates the composition of the series along the period.

This data were not used directly from the acquisition process. Transformations, cleaning and other operations were performed. The details and procedures for the use of this data set are given in [9].

3.2 Data from the lab test

The tests were performed on elastomer elements used in the Flexjoint, to determine the stress/strain properties of the elastomer under various modes of deformation. Quasi-static tests were performed with sequential application of load and unload for increasing levels of deformation. It should be noted that the hysteresis cycle (Figure 3a) appear even under cyclic static loads (quasi-static).
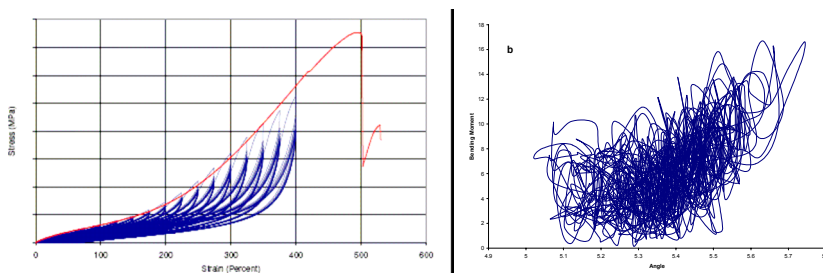

Figure 3:   Hysteresis cycles – a) experimental data b) real data.

4   Selection of parameters of the selected networks

In this part of the study, the concern was to find the most efficient networks to deal with the system dynamics and the hysteresis of the Flexjoint elastometer. Three types of neural network architectures were used to assess the networks' ability to model the hysteretic behavior. In addition to TDNN and recurrent NARX type (PM and SPM) networks, a conventional backpropagation (MLP) network was also modeled. Conventional backpropagation is a static process and does not give good results in processes in which time is a crucial factor. However, it was used in this research only to compare the results and evaluate the gain when time is neglected in systems and materials exhibiting hysteresis.

The traditional gradient-based training algorithms [1], often used in conventional backpropagation networks, can also be used to train dynamic networks. As the dynamic networks processes the information in time order, training with these algorithms can be significantly slower [10]. Also, according to [10], the use of an algorithm based on the optimization method of Levenberg-Marquard provides a more suitable training for the problem.

Training is supervised and stops by the Early-Stopping criterion or if the maximum of 600 epochs (iterations) is reached. In this study, the effectiveness of the models developed was measured by the RMSE (Root Mean Squared Error). To understand the selected networks behavior and to define some parameters, in the following section a series of tests were performed using the experimental data, which presented a volume of data much better behaved than the real data (see Figure 3a and b).

4.1 Analysis and results with experimental data

The laboratories data set is made up of 180 hysteresis cycles, which corresponds to 52194 registers and is stored in sequence, i.e., for a determined arbitrary point of the stress-strain state, the points that preceded it are states which happened previously.

4.1.1 Conventional backpropagation without feedback

The MLP network performance was observed through an available data set sample. Training stopped after 48 epochs by the early stopping criterion with a RMSE = 0,036. Tests were also performed without the validation set and there was no overtraining for the established number of epochs, however, training lasted longer. As was expected, the model with the conventional backpropagation behaves as a nonlinear regression model, as can be seen in Figure 4.

4.1.2 Time Delay Neural Network (TDNN)

The TDNN network input layer is composed of the input variable and its delays, temporally spaced in 1 time unit, t-1, t-2,...,t-N.
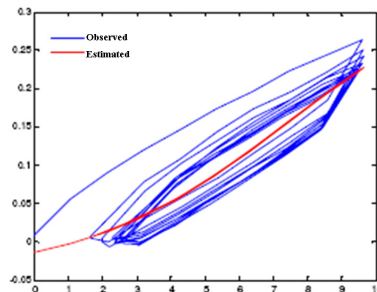

Figure 4:   Stress x Strain – obs and est – conventional backpropagation model.

TDNN approximates a function f given by the equation: $\sigma(t) = f(\varepsilon(t), \varepsilon(t-1), ..., \varepsilon(t-N))$ in which, ε(t) strain in time t ; σ(t) stress in t time and N : sliding window size.

The TDNN training, which works with sequential data [11], is obtained by feeding the network with a fixed window size on the data inputs, obtaining an output and properly adjusting the network weights. This process is repeated for more than one window, placed one step ahead, until all the available data was used.

Usually, the choice of the window size influences the network performance [1]. The methodology adopted to discover the number of optimal delays (lags) was empiric, by means of simulations comparing each model's performance and choosing those with the lowest RMSE.

Different time windows were used to improve the model's performance, considering that is no systematic way to determine how far a system must look back in order to predict the current value [12].

Training stopped by the early stopping criterion with 68 epochs and RMSE(training,test)= (0.1163, 0.2355) . Cyclic systems, like the ones exhibiting hysteresis, are not well represented only by the past of their inputs, but by the past input-output order pair (in this case ε-σ).

4.1.3 NARX recurrent networks

As the previous section demonstrated, only the inputs are lagged in TDNN networks. However, for hysteresis modeling it is necessary to use the pair (stress-strain) which makes up the state of the last load as a parameter for the model's input. Thus, the Nonlinear AutoRegressive with eXogenous inputs recurrent network (NARX) meets with this requirement and its applicability will be investigated herein.

NARX networks are autoregressive nonlinear networks, i.e., the output variable depends on its past, which is why it should also be used as the system's input. In addition, external variables (exogenous) are also part of the model's input data. There are two types of architecture: parallel and series-parallel. Some authors consider NARX as series-parallel and NOE (Non-linear Output Error model) for the parallel model.

NARX parallel-type networks are those in which the model's output are fedback with time lag input variables, together with external input variables. On the other hand, the SPM networks use the output history, i.e., the actual observed values, to predict future values, instead of the values estimated by the model, in addition to the external input variables.

NARX parallel and SPM networks are described by the following equations, $\hat{\sigma}(t) = f(\varepsilon(t), [\varepsilon(t-1), \hat{\sigma}(t-1), ..., [\varepsilon(t-N), \hat{\sigma}(t-N)])$ and

$\hat{\sigma}(t) = f(\varepsilon(t), [\varepsilon(t-1), \sigma(t-1), ..., [\varepsilon(t-N), \sigma(t-N)])$ respectively, where $\hat{\sigma}(t)$ to the approximation of $\sigma(t)$.

To determine the number of time lags and the number of neurons of the hidden layer, the same procedures used for the TDNN networks were adopted.

The analysis concluded the same memory size, 5 lags and the same number of neurons: 10.

Recurrent SPM present a better performance and stability since they use the observed output data as input, not including the error of the output data estimated by the neural network. The best results simulation resulted in RMSE(0.0081, 0.0098) for SPM and RMSE(0.0865, 0.1154) for PM . A more detailed discussion about the convergence of the PM can be found in [2] and an example can be found in [13]. However, the PM approximates the system dynamic, which means, it is able to find out the fundamental laws that regulates the system behavior, while the SPMs "find out" the regularities or periodicities dominant in the [14]. The second aspect concerns the cost of output data acquisition when using the SPM in production. For the networks in production, the choice between PM and SPM networks is crucial because the use of the SPM network requires that the output data must be monitored to be used as input. This cost does not exist when the PM network is used. For training purposes, in both cases the availability of the output data will be necessary, since both models need this data to calculate the error.

Another way to compare is through training time. The PM require much more time than the SPM networks. For example, to train the SPM model with all cycles it took approximately 10 minutes; whereas in the PM 3600 minutes were spent (both cases ran under the same hardware and software configurations).

5   Analysis and results based on monitored real data

Although laboratory data has no expressive dynamic effect because they have been obtained by tests using quasi-static cyclic loading, the strategies used to represent the hysteresis exhibited by the material could be the same to represent the dynamic of the system due to the nature of loading. In both cases, the more appropriate ANN are those representing memory (or time) through the inclusion of time delays in their structure. Therefore, the methodology used in item 4.1.3 serves for the application of data obtained from the real system in operation, where the dynamic and hysteresis are present.

5.1 Data for training/validation/test

Laboratory data refers to the stress and strain, and data available on the real system correspond to the force and displacements (rotation). This study adopted as input the bending moment (kN.m) and tension (kN), and as output the displacement (rotation) in degrees, only as a question of more common physical logic, the forces as excitements, and displacements as responses. Tests have proved that the relation is reversible, that is, the displacements can be either input or output.

As described in section 3, data is composed by sequences in time, each sequence representing one sea state, current and wind, separated from each other by a constant period of time corresponding to 2.5 hours. Due to memory limitation, only 11% of the data was selected randomly from the period, with

due care to maintain the representation of the Flexjoint behavior with series for training, validation and testing.

It is important to observe that each series is understood as a single behavior because it represents a sea state that is, of course, associated with the following period (since the sea is the same place in a further moment), but from the statistics point of view is independent. Because of this, it is not a problem for the system solution the presentation of the series model in random order, that is, that are not in the order they were observed.

Another relevant aspect is the memory representation in the model. Each sequence represents 30 minutes every 3 hours. And, as dynamic networks work with time delays, the beginning of a new sequence does not consider the past of the end of the previous sequence, once the loading conditions are theoretically different after 2.5 hours. This problem would bring some decrease in the network performance due to the noise present at the beginning of each sequence. But, apart from this disturbance occurring only at the beginning, the sequences are long enough to dissipate this effect. This fact has been proved by the analysis carried out. Even when this effect dissipates, it was decided to eliminate the first records (number of records eliminated is a function of the lag adopted in the model) of each series, being presented to the model only as past for the first records duly presented to the neural network.

5.2 Comparison between series-parallel and parallel models

In this section, the results of the nonlinear dynamic system for the Flexjoint are presented through the models selected in the section 4, NARX SPM and PM.

Network training parameters were: 24 Units at the input layer; 10 units at the hidden layer; 2 units at the outlet layer; 104 training patterns (series); 600 epochs; activation function: Sigmoid; number of time delays (sliding window) equal to 5; training algorithm: Levenberg-Marquard; overtraining criterion; Early stopping; MATLAB version 2006Ra (newarxsp and newnarx functions)

The PM yielded RMSE(training,test)=(0.1621, 0.3803) for Rotation 0^0 and RMSE= (0.2652, 0.3663) for Rotation 90^0. In other hand, the SPM yielded more accurate results with RMSE= (0.0238, 0.0568) for Rotation 0^0 and RMSE= (0.0289, 0.0516) for Rotation 90^0.

Results obtained for the training and test sets are shown in Figures 5 and 6. These results represent the best results obtained for each structure after several runs, by modifying the parameters associated with each structure as learning rates, initial weight values, time delays and training set, and considering the best performance for each model.

5.3 Exogenous variables influence on the models

Some tests were carried out to assess the exogenous variable influence (bending and tension) in the model behavior through reanalysis by using its own neural network, by removing all exogenous variables of the model and obtaining its RMSE performance. For Rotation 0°, the simulation resulted in R=(0.0254, 0.0453). The results for Rotation 90° were R(0.0266, 0.0413).

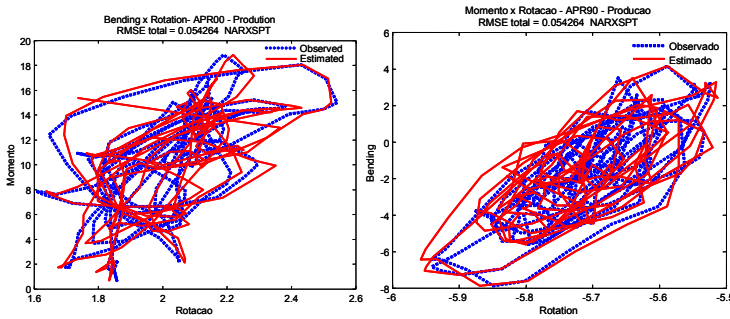Figure 5:   Bending x Rotation – observed/estimated (0º e 90º) – SPM.
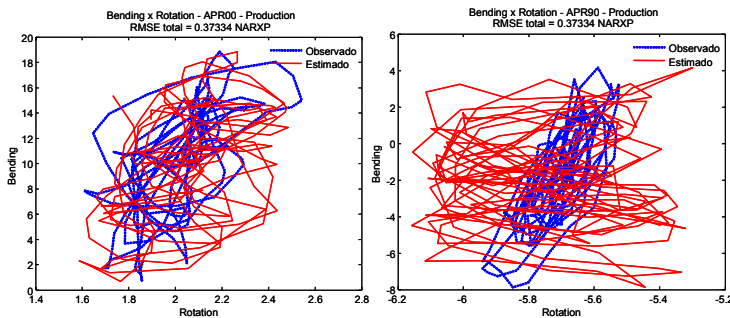

Figure 6:   Bending x Rotation – observed/estimated (0º e 90º) –PM.

As can be observed, there is little influence from the exogenous variables. This result is surprising due to the knowledge of physical phenomenon and good correlation between the exogenous and output variables.

One of the assumptions used to explain the weak contribution of exogenous variables in the SPM was that in which the sequences could have been sampled with a high acquisition rate. To reduce the number of sequence points without jeopardizing the signal quality, Shannon's proposition was adopted ensuring that if Niquist criterion [10] is satisfied, then the signal may be reconstructed with no loss of information. Before decimation, however, a low-pass filter was used in order to eliminate high frequencies. The main reason for this filtering is because the equipment used for the measurement of top inclinations is not able to properly capture the high frequencies. Consequently, it was understood that mapping forces and rotation would be better represented in a wave frequency band, that is, in low frequencies [9].

New analyses were carried out, and the SPM with exogenous variables and those with filter (SP-CF) showed better results than those with no filter (SP_SF). Test RMSEs are 0.0243 and 0.0568 respectively for output variable at 90º angle. While those models with filter (SF_CFD) and decimation (RMSE=0.0560) do not show significant differences compared to models with no filter (SF_SF) (RMSE=0.0568). One of the assumptions raised is the loss of signal quality as a

result of decimation, in spite of the proven gain with the filtering. All models with no output variables show a slightly better performance than those with the output variables.

Performances of the PM, trained and tested in parallel are significantly lower than the SPMs. Here also the decimation seems to have prejudiced the model performance. For an exogenous variable at 0º angle the RMSEs for P_CFD and P_SF models are 0.6688 and 0.3803 respectively. Considering that the training errors were small in the SPM, this can be converted in parallel, and used for the long term predictions. The SPM used in this study are for "one-step-ahead" predictions. Further the models SP_P_CFD (SPM converted in PM with Filtering and Decimation) and SP_P_SF (without Filtering) were generated.

The results were very promising. As shown in Figure 7, despite the discrepancies, the model can reasonably follow the system behavior up to 2000 seconds (30 minutes) for an output variable at 0º angle.

A possible explanation for such a difference in performance may be based on the importance of the variables. Output variable APR0 (angle at 0º direction) has a correlation with the exogenous variables (NUM0 and NUT) much higher than the APR90 variable (angle at 90º direction) with their correlatives (NUM90 and NUT).

The results of the models trained in SPM and tested as PM without filter, are lower than the models with filter and decimation. In general, the output variable APR0 has a better performance than the APR90.
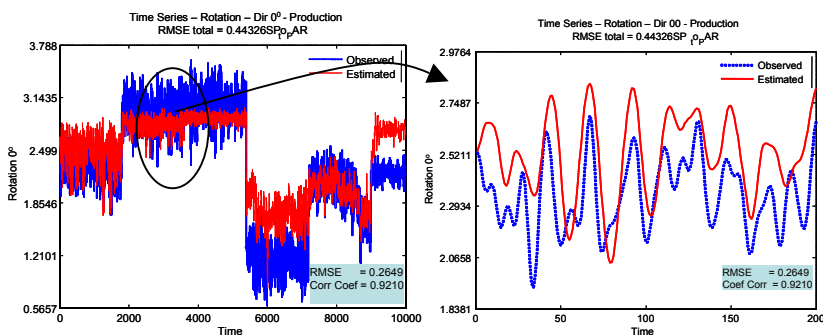

Figure 7:   Model SP_P_CFD – observed versus estimated.

Trained networks through the SPM architecture, and converted into parallel for the production showed quite promising results. The increase in quality of variables may significantly enhance this type of approach.

Another important aspect provided by the analyses results was the investigation on the influence of exogenous variables. Both the laboratory data, or the real data obtained from the monitoring process presented comparable behavior. It is possible that this problem has this particularity, that is, the response does not depend upon exogenous variables. The response has sufficient information (memory) to predict its own future behavior.

6 Conclusions

A data-based model was developed using the ANN technique in order to reproduce the hysteretic behavior of an elastomeric top joint installed on a rigid importation gas riser on Petrobras' PXVIII platform.

Motivation for the work is to obtain a wider knowledge of the joint's behavior in order to check the joint's original characteristics are affected reflecting stress at the riser top region.

The work was based on riser measurements carried out from 1998 to 2002 in order to obtain validations of methodology and assumptions for the riser project. At the time of monitoring system planning there was no intention to use the data for the generation of an empiric model system to simulate the physical properties of equipment.

An ideal situation would be to use laboratory tests under controlled conditions to provide an initial calibration of the model, which would allow the possible to filter erroneous data, or even field equipment errors when adjusting the model with real data. This procedure would also serve as a means of making a critical analysis of the field data obtained. Despite the difficulties, the huge amount of data and complexity of the problem, a pioneer initiative in the riser project and development was taken to characterize the behavior of the top joint, seeking to characterize the top joint behavior, investigating the various possible parameter correlations that are not considered in the traditional equipment design.

The results suggest that this model is rate independent which allowed to conclude that the elastomer works relatively limited stress levels. This information is relevant since the high levels of deformation could precipitate a deformation process of the elastomer [15].

Another relevant result of the joint behavior was not conclusive in regards to its dependency on tension oscillation. For the riser project the independence hypothesis is assumed. The investigation carried out covered high frequency issues, that is to say the dynamic response range associated with vibrations introduced by vortex.

The model response, when eliminating the high frequencies (above 0.25 Hz), showed a significantly improved adjustment. When investigating the causes of this, together with the monitoring technicians the conclusion was that the top inclination measurement equipment was not capable of adequately capturing high frequencies. Based on this conclusion a new specification for future inclinations measurement equipment is underway.

Due to the above problems the doubt remains regarding if the joint response standard is subject to change by the high frequency vibrations to which the riser is subjected.

In addition to bringing relevant information to riser projects and monitoring systems, the proposed model will be used for future rigid riser top monitoring studies.

References

[1] Haykin, Simon / Redes Neurais - Princípios e pratica; 2° edição; editora Bookman, 2004.

[2] Narendra, K. & Parthasarathy, K. Identification and Control of Dynamical Systems using Neural Networks, IEEE Transactions on Neural Networks 1(1): 4–27, (1990).

[3] Elman, J. L. (1990), Finding structure in time. Cognitive Science, 14(2): 179–211.

[4] Basheer A., P.E., M. Asce, Stress-Strain Behavior of Geomaterials in Loading Reversal Simulated by Time-Delay Neural Networks, ASCE - Journal of Materials in Civil Engineering, Vol. 14 No. 3 May–June 2002.

[5] G. W. Ellis, C. Yao, R. Zhao, and D. Penumadu, "Stress-strain modeling of sands using artificial neural networks." International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts, Volume 33, Number 3, April 1996, pp. 115A–115A(1).

[6] Sudheer, K. P., Jain, S. K., Radial Basis Function Neural Network for Modeling Rating Curves, J. Hydrologic Engrg., Volume 8, Issue 3, pp. 161–164 (May/June 2003).

[7] Tawfik M, Ibrahim A, Fahmy H. Hysteresis sensitive neural network for modeling rating curve. ASCE J Comput Civil Engng 1997; 11(3): 206–211.

[8] Machado, R.Z.; Mourelle, M.M.; Franciss, R.; Silva, R.M.; Lima, C.S.; Eisemberg, R.; Oliveira, D., "Monitoring Program for the First Steel Catenary Riser Installed in a Moored Floating Platform in Deep Water", OCEANS'99 MTS/IEEE. Riding the Crest into the 21st Century Volume 2, 1999: 801–810.

[9] Rocha, Djalene Maria, Redes Neurais para Modelagem de Sistemas Estruturais Offshore Dinâmicos não Lineares com Histerese, (COPPE/UFRJ, D.Sc., Engenharia Civil, 2007).

[10] MATLAB HELP Software, version R2006A.

[11] Vojtko, J., Neural Network, Component of Measuring Set for Error Reduction, Measurement Science Review, Volume 4, Section 1, 2004.

[12] T. Z. Tan, C. Quek, G. S. Ng, "Brain-inspired genetic complementary learning for stock market prediction", IEEE Congress of Evolutionary Computation, Vol. 3, Edinburgh, Scotland, 2–5 September 2005, pp. 2653–2660.

[13] Calôba, L. P. Livro de Minicursos do XIV Congresso Brasileiro de Automática, Natal, 2002.Introdução ao Uso de Redes Neurais na Modelagem de Sistemas Dinâmicos e Séries Temporais. COPPE & EP – UFRJ (caloba@ufrj.br).

[14] Von Zuben, F.J, "Modelos Paramétricos e Não-Paramétricos de Redes Neurais Artificiais e Aplicações", Faculdade de Engenharia Elétrica e de Computação (FEEC) - Unicamp, SP, Brasil, 1996.

[15] Burtscher, S., Dorfman, A., Bergmeister, K., 1998. Mechanical aspects of high damping rubber. 2nd. Int. PhD Symposium in Civil Engineering, Budapest.