

MIT Open Access Articles

Using Machine Learning to Identify Important Parameters for Flow-Induced Vibration

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Ma, Leixin, Resvanis, Themistocles L and Vandiver, J Kim. 2020. "Using Machine Learning to Identify Important Parameters for Flow-Induced Vibration." Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE, 4.

As Published: 10.1115/OMAE2020-18325

Publisher: ASME International

Persistent URL: <https://hdl.handle.net/1721.1/139743>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



USING MACHINE LEARNING TO IDENTIFY IMPORTANT PARAMETERS FOR FLOW-INDUCED VIBRATION

Leixin Ma

Massachusetts Institute of Technology
Cambridge, MA, USA

Prof. J. Kim Vandiver

Massachusetts Institute of Technology
Cambridge, MA, USA

Themistocles L. Resvanis

Massachusetts Institute of Technology
Cambridge, MA, USA

ABSTRACT

Vortex-induced vibration (VIV) of long flexible cylinders in deep water involves a large number of physical variables, such as Strouhal number, Reynolds number, mass ratio, damping parameter etc. Among all the variables, it is essential to identify the most important parameters for robust VIV response prediction. In this paper, machine learning techniques were applied to iteratively reduce the dimension of VIV related parameters. The crossflow vibration amplitude was chosen as the prediction target. A neural network was used to build nonlinear mappings between a set of up to seventeen input parameters and the predicted crossflow vibration amplitude. The data used in this study came from 38-meter-long bare cylinders of 30 and 80 mm diameters, which were tested in uniform and sheared flows at Marintek in 2011. A baseline prediction using the full set of seventeen parameters gave a prediction error of 12%. The objective was then to determine the minimum number of input parameters that would yield approximately the same level of prediction accuracy as the baseline prediction. Feature selection techniques including both forward greedy feature selection and combinatorial search were implemented in a neural network model with two hidden layers. A prediction error of 13% was achieved using only six of the original seventeen input parameters. The results provide insight as to those parameters which are truly important in the prediction of the VIV of flexible cylinders. It was also shown that the coupling between inline and crossflow vibration has significant influence. It was also confirmed that Reynolds number and the damping parameter, c^* , are important for predicting the crossflow response amplitude of long flexible cylinders. While shear parameter was not helpful for response amplitude prediction.

Keywords: Flexible riser; Flow induced vibration; Feature selection; Machine learning

Nomenclature	
$A_{rms,cf}$	Spatial and temporal RMS of cross-flow response amplitude in the power-in region
$A_{rms,il}$	Spatial and temporal RMS of in-line response amplitude in the power-in region
A_{cf}^*	Dimensionless cross-flow amplitude for flexible cylinders in the power-in region
A_{il}^*	Dimensionless in-line amplitude for flexible cylinders in the power-in region
D	Cylinder diameter
$EI(z)$	Cylinder's bending stiffness
L_{in}	Length of power-in regions
L	Cylinder length
P_0	Initial tension
P	Dynamic mean tension during VIV process
$U(z)$	Flow speed
U_{rms}^2	Mean square flow velocity in the excitation region
$\Delta U / L$	Shear gradient of the current profile
$\bar{V}r$	Reduced velocity
$c(z)$	Damping coefficient, damping/length
$c_{e,cf}$	Cross-flow equivalent damping coefficient
$c_{e,il}$	In-line equivalent damping coefficient
$c_h(z)$	Hydrodynamic damping coefficient
$c_s(z)$	Structural damping coefficient
c^*	Dimensionless damping parameter for rigid cylinders or flexible infinite cylinders
k_{cf}	Cross-flow wave number

k_{il}	In-line wave number
$m(z)$	Cylinder's mass/length in air
$x(z, t)$	VIV in-line displacement
$y(z, t)$	VIV cross-flow displacement
z	Axial coordinate along the pipe
α_{cf}	Cross-flow travelling wave index
α_{il}	In-line travelling wave index
β	Shear parameter
κ_{cf}	Cross-flow mode dominance factor
κ_{il}	In-line mode dominance factor
ρ	Density of fluid
μ	Dynamic viscosity of fluid
ω_{cf}	Cross-flow response frequency in radians / second
ω_{il}	In-line response frequency in radians / second
ζ	Structural damping ratio

INTRODUCTION

Vortex-induced vibration (VIV) is a common example of fluid structure interaction, which is known to be influenced by a wide variety of dimensionless parameters, including the Strouhal number, Reynolds number, mass ratio, and damping. (Vandiver 1993, 2012; Sarpkaya 2004).

To predict VIV response amplitude, many dimensionless parameters have been proposed over the past 65 years. Scruton first introduced a form of the well-known damping parameter, $m^*\zeta$ in 1955, where m^* and ζ are the mass ratio and structural damping ratio, respectively. High correlation between $m^*\zeta$ and peak amplitude A^* was observed for single mode response of cantilevers' and spring-mounted rigid cylinders. When Griffin tried to extend the use of $m^*\zeta$ to predict the amplitude response of all types of cylinders, including flexible ones, a great deal of scatter was left unexplained, (Griffin, 1975). In 2006 Govardhan and Williamson showed that when Reynolds number was taken into account, the peak response of spring-mounted, rigid cylinders could be predicted with very little variance, (Govardhan & Williamson, 2006).

It was not until 2018 that a dimensionless damping parameter, c^* , was proposed, which clarified the role of damping in determining the response of flexible cylinders, especially in complex sheared flows, (Vandiver et al, 2018).

There remain dozens of possible dimensionless parameters that might influence the response of long flexible cylinders in sheared flows. Field experiments, laboratory experiments and direct numerical simulations have revealed that when long tensioned cylinders are exposed to spatially non-uniform current profiles, they can exhibit a mixture of standing and traveling waves in both the in-line and cross-flow directions (Bourguet 2011; Vandiver et.al. 2018). Moreover, the VIV is often not stationary with chaotic and stochastic response often observed, (Modarres-Sadeghi, 2011). The objective of this research was to

determine the minimum number of dimensionless parameters necessary to predict VIV response with reasonable prediction error.

In AI terminology, a feature is equivalent to a dimensionless parameter, which is used to predict a desired outcome. Thus we seek the minimum number of features needed to adequately predict the desired response. Techniques are needed to discriminate between those features that are truly important and those that may be left out. Due to the complicated correlations of the VIV features, the accumulated experience that comes from the work of Scruton, Sarpkaya, Govardhan, Williamson and others can be treated as prior knowledge to constrain the AI optimization and ensure that the desired important parameter set with the prior knowledge is included.

In this study, three parameters were selected to be included in this prior knowledge group: Reynolds number, the c^* damping parameter, and the shear parameter, β , which was thought to be needed to help distinguish between uniform and sheared flows. It is good practice to test any assumptions one makes in setting up a model. This is done at the end of the study to test whether or not the three prior knowledge selections were, in fact, valid. It is revealed later that one of the three was not in fact justified for this dataset.

The measure of performance of a selected set of features is the prediction accuracy. Various feature selection methods have been proposed with the goal of assessing the features' roles in improving the prediction accuracy in a machine learning model (Guyon, 2007). Since the machine-learning based approach can take into account complicated interactions among features, this approach will be applied to identify important parameters (or features) for vortex induced vibration.

In this paper, $A_{rms,cf}/D$, the flexible cylinder's cross-flow, dimensionless VIV amplitude in the power-in region was chosen as the prediction target, or dependent variable. A Deep Neural Network (DNN) approach was used to build nonlinear mappings between seventeen dimensionless input parameters and the crossflow vibration amplitude. The dataset used in the study came from 38-meter-long cylinders with diameters of 30mm and 80 mm, which were tested in uniform and sheared flows at Marintek in 2011. When all seventeen parameters were included a minimum prediction error of 12% was realized. By sequentially selecting the features that gave the greatest prediction accuracy improvement on top of the three base parameters that form the prior knowledge, we were able to achieve a prediction error of only 13%, using just six of the original seventeen input parameters. In the next few sections the methodology for selecting important parameters and carrying out the machine learning procedure is presented. It was found that the coupling between in-line and cross-flow vibration can have a significant influence on cross-flow response amplitude prediction for certain cases. Finally, the machine learning predictions confirmed that Reynolds number and the damping parameter, c^* , are important for predicting long flexible cylinder's cross-flow VIV amplitude. While differentiating the

sheared flow and uniform data using a shear parameter was not necessary for improving response amplitude prediction. This could be because the damping parameter was sufficiently different for the uniform and sheared flow tests and in this case can differentiate the data.

STRATEGY DEFINITIONS AND OBJECTIVES

Blum (1997) gave a formal definition for a feature's incremental importance: Given a sample of data S , a learning algorithm L , and a feature set X , feature x_i is incrementally useful to L with respect to X , if the accuracy of the hypothesis that L produces using the feature set $\{x_i\} \cup X$ to predict target y is better than the accuracy achieved using just the feature set X . The machine learning algorithm L may be chosen as a deep neural network (DNN), a random forest, or a kernel machine etc.

This definition suggests that by incrementally removing or adding a test feature x_i on top of a certain pre-selected feature set X , a large influence on the prediction accuracy shows that test feature is important to the prediction target. On the other hand, small influence on accuracy suggests that test feature does not contribute in prediction, which may be due to its irrelevance to the prediction target or redundancy with other input features.

For flow-induced-vibration problems, the prediction target is defined here as the cross-flow vibration amplitude. Previous VIV research has revealed the importance of some features such as Reynolds number on vibration amplitude. Given this prior knowledge, the problem objective is to identify the minimum number of additional features that can complement the prior knowledge and achieve a level of prediction accuracy that is close to the full feature set. The full feature set is described below. It contains 17 parameters common to VIV experiments and predictions. The objective is to find the smallest subset of these 17 features that achieves nearly the same prediction error as the full set of 17.

1. The Forward Greedy Approach Strategy

The strategy for identification of important features is conducted in the following manner.

First, two baseline predictions are obtained using two different feature subsets. One uses all 17 features. The second uses only the three features designated as prior knowledge. For each prediction, an error metric is computed averaged over all of the available data. When all 17 candidate features are included in the model, the average prediction error, E_a , is smallest, because the effects of all the input features and their interactions are accounted for in the machine learning model. When only the prior knowledge is included, due to the absence of additional key features, the average prediction error defined as E_p is usually greater than and at best equal to E_a . Then a systematic strategy is implemented to determine those additional features that are most effective in reducing the prediction error. The strategy used here is commonly known as "forward greedy".

To find an optimum subset of features requires conducting a search process using various feature combinations and comparing the machine learning predictions that result from each combination. For m additional parameters, a complete combinatorial search requires comparing 2^m predictions. A

'forward greedy' procedure is more efficient. It is described in Figure 1. In each round of iteration, a new test feature, x_i ($i \in I$) is added on top of the previously selected feature set X_p . The new feature set, $x_i \cup X_p$, is introduced to the machine learning model, and the corresponding prediction error, E , is computed. After comparing the prediction errors E_i ($i \in I$), the test feature that gives the smallest prediction error is chosen and combined with the feature set X_p for the next round of iteration. Such an iterative process continues until the prediction error is no larger than an acceptable error threshold ε (Guyon, 2003).

Finally, the importance of the *a priori* chosen features that make up the prior knowledge can be examined by removing each prior knowledge feature from the identified important feature subset and then comparing the resulting prediction error with that discovered using the forward greedy approach. Any prior knowledge features that end up having a negligible effect on the prediction error can be dropped from the optimum subset.

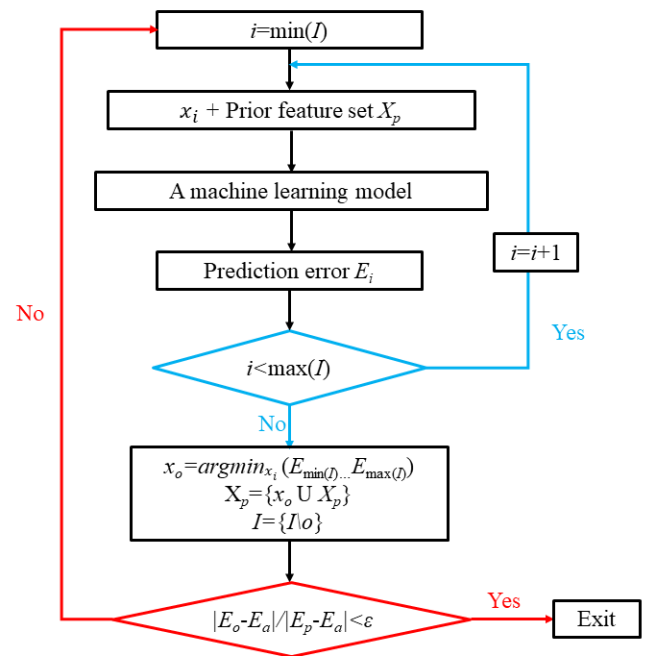


Figure 1 A flow chart for conducting forward greedy feature selection strategy

VIV PARAMETER SELECTION

1. Flexible cylinder VIV modeling

Seventeen dimensionless parameters are chosen to be used to predict the crossflow $A_{rms,cf}/D$ response. The parameter selection began with a study of the equations of motion describing the cylinder's motion in the cross-flow and in-line directions when subjected to VIV. Figure 2 shows a tensioned elastic cylinder under a linearly sheared current profile $U(z)$ distributed along axis z , that is free to vibrate in both the in-line

(IL) and cross-flow (CF) directions. The equations of motion in cross-flow direction and in-line direction can be expressed as,

$$m(z) \frac{\partial^2 y}{\partial t^2} + c_s(z) \frac{\partial y}{\partial t} - P(z, t) \frac{\partial^2 y}{\partial z^2} + EI(z) \frac{\partial^4 y}{\partial z^4} = F_{cf} \quad (1)$$

$$m(z) \frac{\partial^2 x}{\partial t^2} + c_s(z) \frac{\partial x}{\partial t} - P(z, t) \frac{\partial^2 x}{\partial z^2} + EI(z) \frac{\partial^4 x}{\partial z^4} = F_{il} \quad (2)$$

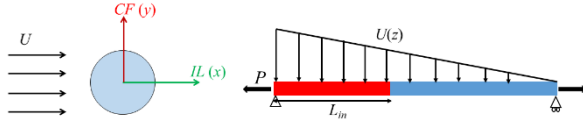


Figure 2 Side and top view of a cylinder under VIV in a linearly sheared current

where $x(t)$ and $y(t)$ are the in-line and cross-flow displacement, respectively. $F_{cf}(t)$ and $F_{il}(t)$ are the vortex induced forces on the cylinder in the two directions, which are the result of complicated interaction between vortex shedding and structural vibration. Under steady-state, narrow-banded vibration, the VIV forces transfer energy from fluid to the vibrating cylinder in a well-defined region with length L_{in} , which is termed as the “power-in” region (Vandiver et.al. 2017, 2018). Outside the “power-in” region, the vortex loading dissipates structural vibration energy through hydrodynamic damping $c_h(z)$. In experiments or numerical simulations, the location of the “power-in” region and the total power dissipation can be identified from structural vibration measurements (Rao 2014). The VIV forces in the “power-in” region may involve at least the following parameters,

$$F_{cf} = f(U(z), x(z, t), y(z, t), \rho, \mu, L, L_{in}, D, m(z), c_{s,cf}(z), c_{h,cf}(z), P(z, t), EI(z)) \quad (3)$$

$$F_{il} = f(U(z), x(z, t), y(z, t), \rho, \mu, L, L_{in}, D, m(z), c_{s,il}(z), c_{h,il}(z), P(z, t), EI(z)) \quad (4)$$

An exact formulation for the VIV forces is not known. Therefore, it is not currently possible to directly solve for the VIV response. However, with machine learning methods it is possible to build quantitative relationships between selected measures of VIV response and other physical parameters, such as Reynolds number, total energy dissipation and reduced velocity. Considering that the VIV response is excited by the vortex shedding in the “power-in” region, an appropriate response metric is the spatial and temporal RMS of the cross-flow response in the power-in region, as defined in equation 5.

$$A_{cf}^* = A_{rms,cf} / D \quad (5)$$

From measured response data it is also possible to estimate additional metrics known to be important in the fluid-structure interaction that determines the non-linear relationships between structural motion and fluid excitation. One factor is the presence of standing or traveling waves. Another is the presence of single or multiple modes in the response. Measures of these parameters

are implemented as features in the machine learning strategies employed in this study. These features are termed global features in this work, because they refer to properties averaged over the power-in region. Additional global parameters are introduced and discussed in the next section.

2. Global VIV features

A typical time series and wavelet analysis for VIV response is shown in Figure 3. To capture the temporal variation for the VIV process, a sliding window approach is employed. For each small temporal window, a complex proper orthogonal decomposition (complex POD) is conducted on the structural displacement in the “power-in” region, such that the displacement is decomposed into a linear combination of orthogonal complex modes (Feeney 2008).

The ratio between the energy of the principal complex POD mode and the total energy is defined as κ , which is a measure of the dominance of the principal mode. When $\kappa = 1$, a single POD mode dominates the response.

Next it is useful to know if that principal POD mode, which is designated by the symbol ϕ , is composed of standing or traveling waves. A parameter, α , is defined which may vary between 0 and 1. For $\alpha=0$ the principal mode, ϕ , is composed of pure standing waves. When $\alpha=1$, it indicates that ϕ is composed of pure traveling waves. ϕ generally has real and imaginary parts. Mathematically, the traveling wave index, α , is defined by the following equation (Feeney 2008):

$$\alpha = 1 / \text{cond}([\text{Re}\{\phi\}, \text{Im}\{\phi\}]) \quad (6)$$

Where, $\text{cond}(\phi)$ is the condition number of ϕ , which is the ratio of the largest singular value of ϕ to the smallest. For instance, Figures 4 and Figure 5 show a typical VIV measurement from the 38-m-long, 30 mm diameter, cylinder in a linearly sheared current. The traveling wave index, α , and the modal dominance factor, κ , are evaluated within a moving window, which is 3 periods of vibration in length. In this example, the response is dominated by travelling waves propagating from the high speed end of the sheared current to low speed end. The computed κ values in the CF direction are close to 1 which means that the VIV response is best described by a single complex POD mode. Note, that this mode varies in time between being a travelling wave and a standing wave as is evident by the variation of the computed value of α , the traveling wave index.

The vibration process in the in-line direction is more non-stationary. For instance, when the cross-flow vibration is single mode dominated, the in-line vibration may require the superposition of several modes and the vibration may be influenced by different combinations of standing waves and travelling waves.

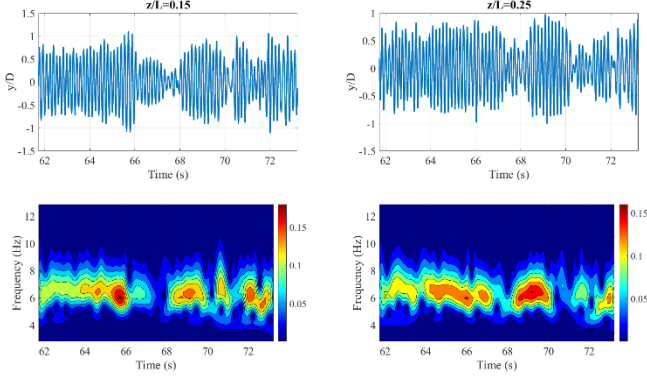


Figure 3 Top: Time series of crossflow displacement at two locations in the power-in region; Bottom: Wavelet analysis on the measured time series. (2011 Shell experiment, $D=30$ mm, linearly sheared flow, $U_{max}=1.6$ m/s)

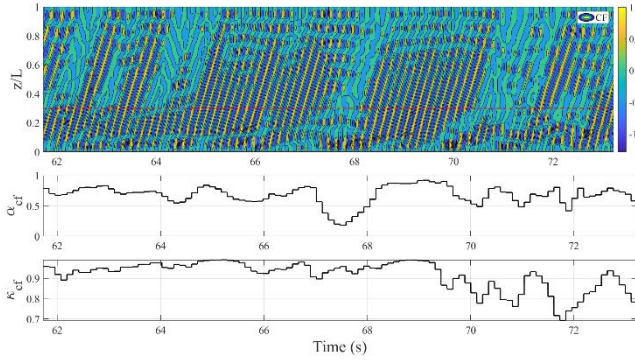


Figure 4 Contour plot of cross-flow vibration around the dominant frequency, its corresponding travelling wave index α_{cf} , and mode dominance factor κ_{cf} in the estimated “power-in” region ($z/L=0\sim0.3$, linearly sheared flow, $U_{max}=1.6$ m/s at $z=0$)

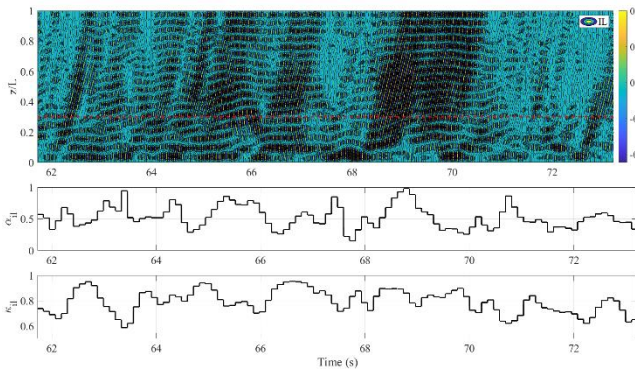


Figure 5 Contour plot of in-line vibration around the dominant frequency, its corresponding travelling wave index α_{il} , and mode dominance factor κ_{il} in the estimated “power-in” region ($z/L=0\sim0.3$, linearly sheared flow, $U_{max}=1.6$ m/s at $z=0$)

The total power dissipation $\langle \Pi_{out,cf} \rangle$ resulting from the distributed hydrodynamic damping and structural damping can be represented by the dissipation from an equivalent rigid cylinder with length equal to the power-in region. This equivalent damping coefficient, $c_{e,cf}$, may be defined in terms of the total dissipated power and the cross-flow mean square velocity $A_{rms,cf}^2 \omega_{cf}^2$ (Vandiver et.al. 2018), as,

$$c_{e,cf} = \langle \Pi_{out,cf} \rangle / (L_{in} A_{rms,cf}^2 \omega_{cf}^2) \quad (7)$$

Hence, for a homogeneous, tensioned cylinder, which exhibits narrow-banded vibration in uniform or linearly sheared current, the authors have chosen to express the response given in equation (5) with an approximate one expressed in terms of the global vibration features, as shown in Equation (8).

$$A_{rms,cf} = f(U_{rms}, \Delta U, A_{rms,il}, \omega_{cf}, \omega_{il}, k_{cf}, k_{il}, \alpha_{cf}, \alpha_{il}, \kappa_{cf}, \kappa_{il}, \rho, \mu, L, L_{in}, D, m, c_s, c_{e,cf}, c_{e,il}, P_0, P, EI) \quad (8)$$

k_{cf} and k_{il} are wave numbers that can be estimated from the dispersion relationship for the tensioned beam.

Non-dimensionalizing Equation (8), the crossflow RMS amplitude A_{cf}^* can be expressed as a function of 17 parameters,

$$A_{cf}^* = f(\text{Re}, \beta, A_{il}^*, V_{r_{cf}}, V_{r_{il}}, Lk_{cf}, \alpha_{cf}, \alpha_{il}, \kappa_{cf}, \kappa_{il}, L/L_{in}, L/D, m^* \zeta, c_{cf}^*, c_{il}^*, P_0 L^2/EI, P/(EI k_{cf}^2)) \quad (9)$$

Where $A_{cf}^* = A_{rms,cf}/D$ and $A_{il}^* = A_{rms,il}/D$ are the dimensionless RMS crossflow and inline response amplitude, respectively. $\text{Re} = \rho U_{rms} D / \mu$ and $\beta = (D/U_{rms})(\Delta U/L)$ are the Reynolds number and shear parameter for the current profile; $m^* \zeta = 4m\zeta / (\pi \rho D^2)$ is the mass damping parameter in the VIV literature, which has traditionally been used to predict the peak response amplitude of both rigid and flexible cylinders. It has been included here to test its relevance to the prediction of the response of flexible cylinders. $V_{r_{cf}} = 2\pi U / (\omega_{cf} D)$ and $V_{r_{il}} = 2\pi U / (\omega_{il} D)$ are the crossflow and inline reduced velocities, respectively; $c_{cf}^* = 2c_{e,cf} \omega_{cf} / \rho U_{rms}^2$ and $c_{il}^* = 2c_{e,il} \omega_{il} / \rho U_{rms}^2$ are the dimensionless equivalent damping parameters in the cross-flow and in-line directions, respectively (Vandiver et.al. 2018).

Equation (10) suggests that the cross-flow VIV amplitude may be governed by the variation of these 17 parameters. However, not all the parameters are necessarily important for accurate cross-flow response amplitude predictions. The methods employed in this paper are intended to reveal those that are important.

MACHINE LEARNING APPLIED TO VIV PARAMETER IDENTIFICATION

1. Dataset description

The dataset is from VIV experiments conducted by Shell Oil Co. in 2011 at Marintek. The test matrix analyzed is shown in

Table 1, which includes two 38-meter-long flexible cylinders with different diameters (30-mm and 80-mm) but of the same bending stiffness. The bare cylinders were tested in uniform and linearly sheared current profiles with the maximum flow speed, U_{max} , ranging from 0.5 m/s to 2.5 m/s. Detailed descriptions of the experiments can be found in Lie (2013) and Rao (2015).

The measured structural damping ratio ζ in the experiment was around 0.5% (Vandiver et.al. 2018). Hence, for the 30-mm and 80-mm cylinders, $m^*\zeta$ was 0.0077 and 0.0057, respectively. For each test case, the VIV response data were analyzed when the cylinder's towing speed reached its steady state. The cross flow reduced velocity (based on the measured response frequency), Vr_{cf} , varies in a narrow range from 6 to 9, and $Vr_{cf}/Vr_{il} \approx 2$. Moreover, the ratio $P/(Elk_{cf}^2)$ was greater than 20, which indicates that in most of the cases, the flexible cylinders were tension dominated (Marcollo et.al. 2018). The VIV features were averaged over the power-in region regions. In the uniform flow, the power-in regions were assumed to cover the whole length of the pipe, while in sheared flow, the assumed power-in region was located at the top 30% of the pipe near the high speed end. Both of these decisions were based on prior analysis which had identified the locations and extent of the power-in regions (Zhibiao 2014). Flexible cylinder VIV at times exhibits a mixture of stationary and non-stationary vibration.

In this study, to capture the temporal variation, the measurements were analyzed within a moving window, which was three cross-flow vibration periods in length and had 75% overlap with the previous window. To differentiate the stationary and non-stationary VIV conditions we studied how key features varied over a time interval nine vibration periods in length, viewed through the lens of a moving window, three periods long. VIV sections were deemed to be stationary when the cross-flow and in-line amplitudes varied less than 20% over 9 periods, and the variation of travelling wave indices were less than 10%. Future investigations will be conducted on the effect of stationarity on the machine learning results.

Each unique run of the neural network described here begins with the input of data to one or more of the input nodes. The maximum number of inputs is seventeen. One such unique run is counted as a data point. Each of the input features is computed from a moving average window, which is three cross-flow vibration periods in length. The constant features in a test condition such as Re and initial tension were repeated for each 3 vibration period. Then, the data samples were classified as stationary or non-stationary according to the features' variation over neighboring 8 windows (9 vibration periods). Of all the data samples from the test matrix of 38 m cylinders, most of them were identified to be non-stationary. Therefore, the data analyzed in this paper were composed of samples that were from 20% stationary and 80% non-stationary VIV time series. The total number of data points used in the machine learning predictions described here was 13,000.

2. Machine learning setup and training

A deep neural network (DNN), which has the universal approximation property, was selected as the learning machine to build a nonlinear mapping between inputs and the target output (Bishop, 2006). A fully connected DNN was constructed with two hidden layers. The target output of the neural network is the cross-flow response amplitude A_{cf}^* . The input layer contains a number of input features selected from the 17 dimensionless parameters on the right-hand-side of Equation (10). Each hidden layer was made up of twenty neurons with a sigmoid activation function.

The output feature A_{cf}^* was normalized to values between 0 to 1. The mean absolute percentage error (MAPE) was chosen as the error metric for prediction, which calculates,

$$Error \triangleq MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (10)$$

where y_i and \hat{y}_i are the measured and predicted target output, respectively. n is the total number of data samples. 70% of the data were used as the training data, while the rest were used as the test data.

To prevent the neural network from overfitting the training data, a well-known strategy is to add an additional regularization function to the error function (Bishop, 2006). In this paper, the commonly used Tikhonov regularization was applied, whose corresponding regularized error function (loss function) is expressed as,

$$L = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| + \gamma \|w\|_2^2 \quad (11)$$

Where γ is known as weight decay. $\|w\|_2^2$ is the $L2$ norm of w , where w represents the weights connecting the neurons.

The neural network optimization was conducted to minimize the regularized error of the training data via the mini-batch stochastic gradient descent algorithm. During neural network training, the batch size was 128 and learning rate was set as 0.05. A total number of 6000 iterations were conducted to optimize the unknown weights, w , connecting the neurons. When $\gamma=0.1$, the MAPE for the testing and training data are almost identical, which suggests no overfitting in the machine learning model.

3. Prior knowledge for cylinder's cross-flow VIV amplitude prediction

Analysis of VIV experiments on spring-mounted rigid cylinders show that the cross-flow VIV amplitude increases with increasing Reynolds number in the range from 10^3 to 10^4 and decreases under larger damping parameter c_{cf}^* (Govardhan & Williamson 2006, Vandiver 2012).

Similarly, cross-flow response amplitude of long flexible cylinders was also found to be influenced by the Reynolds

number and the dimensionless damping (Swithenbank 2008, Resvanis 2012, Rao 2015, Vandiver 2018). However, the scatter in the identified (or fitted) relationships suggest that other parameters are also influencing the resulting cross-flow amplitudes.

Since Reynolds number, Re , and dimensionless damping parameter, c_{cf}^* are known to be important, these two parameters are treated as the prior knowledge for cross-flow response prediction. Moreover, to differentiate between uniform and sheared flows, the shear parameter β was also added to the list of prior knowledge. For typical VIV predictions, Reynolds number Re and dimensionless damping parameter c_{cf}^* and shear parameter β can often be easily estimated.

RESULTS

1. Baseline predictions

First, the full set of 17 parameters are input to the DNN, which gives an MAPE of $E_a=12.0\%$ for all the test data, as shown in Figure 6. The MAPE for sheared flow and uniform flow are 11.0% and 13.5%, respectively. While Figure 7 shows that a model using solely the prior knowledge of Re , c_{cf}^* and β leads to an MAPE of $E_p=24.0\%$. This suggests that some additional features are needed to improve prediction accuracy on top of the 3 parameters that form the prior knowledge.

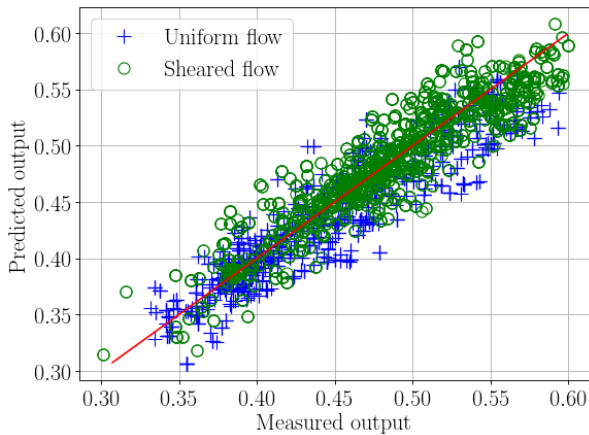


Figure 6 Comparison between DNN predicted and measured cross-flow amplitude using full feature set (MAPE: 12 %)

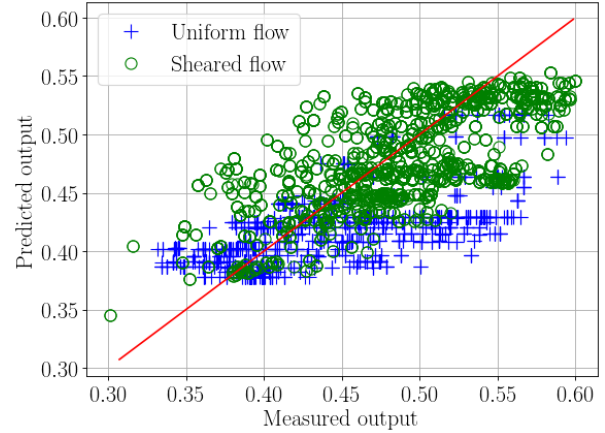


Figure 7 Comparison between DNN predicted and measured cross-flow amplitude using prior knowledge only (MAPE: 24%)

2. Discovery of additional features

Next, on top of the 3 features that form the prior knowledge, the influence of adding each of the other 14 features to the DNN is assessed. Some comparisons between DNN prediction and experimental measurement are shown in Figure 8 to Figure 10.

$m^*\zeta$ is a parameter that accounts for structural damping. For the 30 mm cylinders and 80 mm cylinders, the $m^*\zeta$ were 0.0057 and 0.0077, respectively. Therefore, the introduction of the parameter is able to separate the data to two subcategories of cylinders. Figure 8 shows predictions after the $m^*\zeta$ parameter has been added to the prior knowledge subset. This prediction error improved to just 23.6% from 24%. This suggests that the $m^*\zeta$ has very little effect on the prediction accuracy. This should come as no surprise. $m^*\zeta$ has only two numerical values for the entire uniform and sheared flow datasets. It only accounts for structural damping and not the much greater hydrodynamic damping that arises in a sheared flow. The negligibly small benefit that was observed in the prediction accuracy is most likely due to the fact that the binary nature of this parameter helped to separate the data into two sub-categories for the high-mode number cases of the 30mm cylinder and the low mode number cases of the large diameter cylinder, but it did little to aid in the prediction of RMS response amplitude.

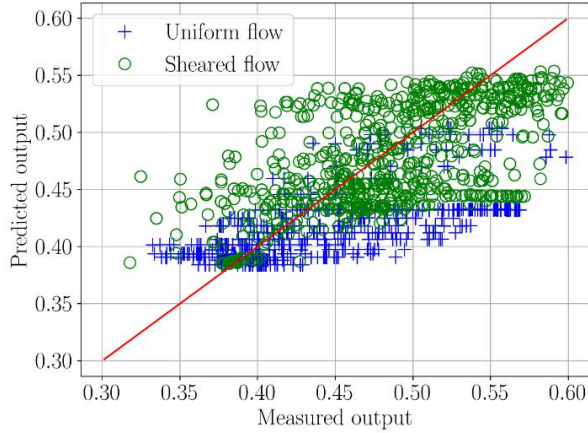


Figure 8 Comparison between DNN predicted and measured cross-flow amplitude using prior knowledge and $m^*\zeta$ (i.e. prediction model using Re , c_{cf}^* , β and $m^*\zeta$) (MAPE: 23.6 %)

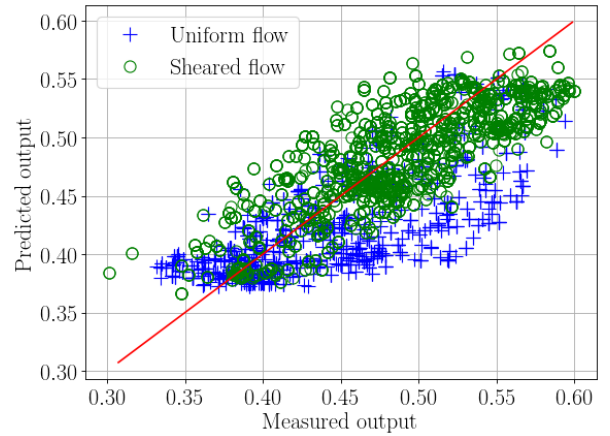


Figure 10 Comparison between DNN predicted and measured cross-flow amplitude using prior knowledge and Vr_{cf} (i.e. prediction model using Re , c_{cf}^* , β and Vr_{cf}) (MAPE: 21.4 %)

Figure 9 to Figure 10 show the effect of adding the in-line response A_{il}^* or the cross-flow reduced velocity Vr_{cf} to prediction model in addition to the prior knowledge. Both of which can help to decrease the mean prediction error. It is known that the reduced velocity plays an important role in rigid cylinders' VIV (Vandiver, 2012). However, the introduction of Vr_{cf} alone does not significantly improve the prediction error. This may be due to the fact that flexible cylinders vibrate within a very narrow range of favorable reduced velocity because the response will shift to another more favorable mode rather than move to an unfavorable reduced velocity. Furthermore, another reason for the small prediction error improvement may be because this range of favorable reduced velocities was found to be strongly correlated with the variation of the Reynolds number which is part of the prior knowledge (Resvanis 2012). Hence the addition of reduced velocity adds little additional information.

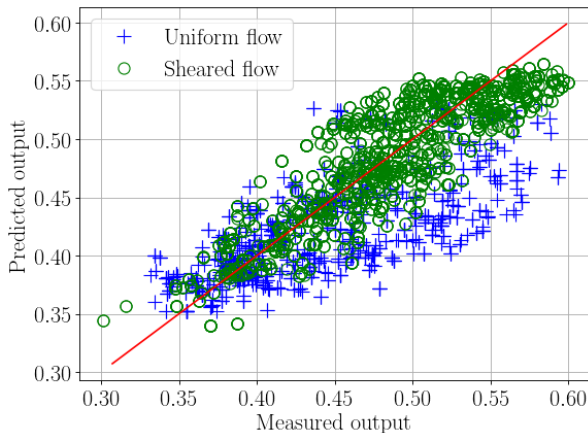


Figure 9 Comparison between DNN predicted and measured cross-flow amplitude using prior knowledge and A_{il}^* (i.e. prediction model using Re , c_{cf}^* , β and A_{il}^*) (MAPE: 18.0%)

In addition to the forward-greedy feature selection algorithm (as shown in Figure 1), we also conducted combinatorial searches for 1 and 2 additional features, by comparing the prediction errors from all possible combinations of the remaining 14 parameters. Adding features using this brute-force approach allows us to investigate the specific effect of each additional feature. Table 2 shows the MAPE for the combinations of additional features on top of the prior knowledge that led to the smallest prediction errors. Δ MAPE is defined as the error reduction caused by the m^{th} additional feature from the smallest error from the prior knowledge and $m-1$ additional features. Adding the inline response amplitude as an additional feature on top of the prior knowledge parameters led to the most significant error improvement of any single additional parameter. What is particularly interesting is that if we focus solely on the sheared flow cases, we see that adding the inline response amplitude reduces the MAPE from 21.2% to 14.5%, which is surprisingly close to the 11% for the sheared flow cases when using all 17 parameters. The prediction error's reduction for the uniform flow cases is not as strongly affected by the addition of just the inline response amplitude, showing just a small improvement from 28% to 23.4%.

In order to observe a significant reduction of the prediction error for the uniform flow cases we need to add the mode dominance factor κ_{cf} or κ_{il} to the prior knowledge and A_{il}^* . Doing so reduces the MAPE in uniform flow decreases from 23.4% to 16.6% but has little effect on the already improved sheared flow predictions. This suggests that in uniform flow cases, under a particular c_{cf}^* and Re , the variation of mode dominance in the VIV response has a big effect on the relationship between in-line amplitude A_{il}^* and cross-flow response amplitude A_{cf}^* . The predictions given the 5-feature subsets (c_{cf}^* , Re , β , A_{il}^* , κ_{il} and c_{cf}^* , Re , β , A_{il}^* , κ_{cf}) both of which

lead to very similar MAPE are compared with experimental measurements in Figure 11 and Figure 12.

Figure 13 shows the variation of minimum MAPE with the number of additional features on top of the prior knowledge using both combinatorial search and forward greedy feature selection. A complete combinatorial search of 3 and 4 parameters out of 14 parameters requires 1092 and 24024 rounds of machine learning predictions, respectively. To reduce the number of predictions, in this paper, the combinatorial search is conducted with other features on top of c_{cf}^* , Re , β , A_{il}^* , κ_{il} . The red symbols in Figure 13 are the predicted results from the combinatorial searches. In each round of combinatorial search, it was observed that the introduction of L/L_{in} , L/D , $m^*\zeta$ had little effect on the prediction. In this dataset the shear parameter β is strongly related to L/L_{in} and the cylinder with larger diameter had the larger Reynolds numbers therefore the, L/L_{in} , L/D , $m^*\zeta$ parameters are redundant with the pre-selected prior knowledge parameters.

When forward greedy algorithm is implemented, inside each iteration, we add each of the remaining features onto the preselected feature set one at a time. The feature that leads to the maximum error reduction is retained in the preselected feature set for next round of iteration. The connected blue circles in Figure 13 are the minimum MAPE in each iteration of the forward greedy selection. When the number of additional feature increases from 1 to 4, the forward greedy selection only requires 50 rounds of neural network predictions, which is much smaller than thousands of neural network predictions required for the complete combinatorial search. Hence, forward greedy algorithm is very well suited to choosing additional features on top of what forms the selected prior knowledge in order to give the smallest prediction error without requiring running thousands of combinatorial predictions

A total of five parameters, including the three from prior knowledge, accounts for most of the reduction in the prediction error. Further introducing two additional features decreases the error from 13.8% to 12.4% for the sheared flow cases. More than one combination of features led to similar results when a total of 7 features were used. This is primarily due to the fact that all the features varied together in the VIV process. When we include more than 4 additional features (i.e. 3 prior and 4 additional for a total of 7), the prediction error is only slightly reduced and asymptotically approaches the error obtained using all 17 features. For example, the prediction using prior knowledge and A_{il}^* , κ_{il} , Lk_{cf} , α_{cf} results in a MAPE of 13% for the combined uniform and sheared data, which is close to the MAPE of 12%, when using all the 17 parameters.

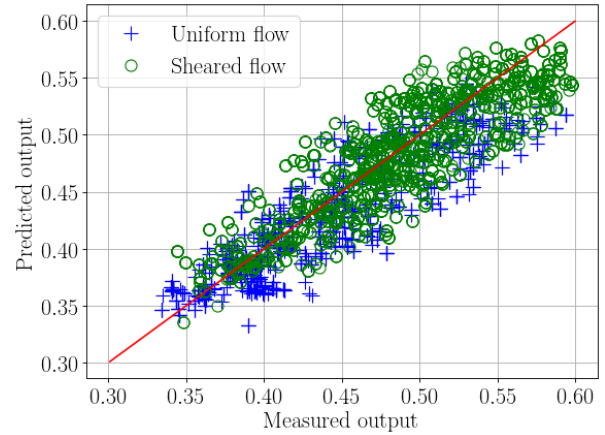


Figure 11 Comparison between DNN predicted and measured cross-flow amplitude using prior knowledge and A_{il}^* , κ_{il} (MAPE: 14.8 %)

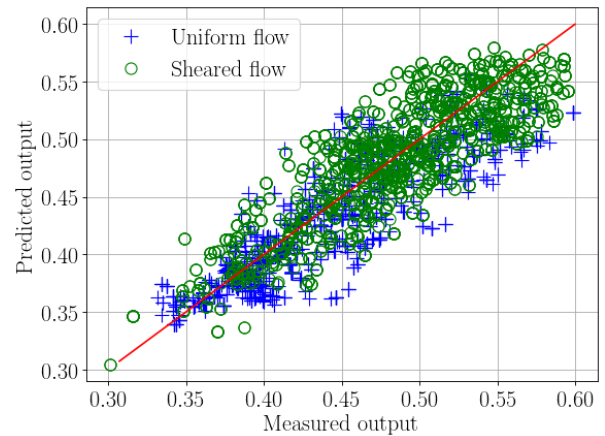


Figure 12 Comparison between DNN predicted and measured cross-flow amplitude using prior knowledge and A_{il}^* , κ_{cf} (MAPE: 15.2 %)

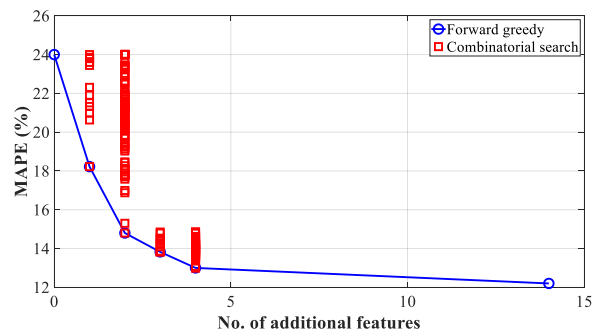


Figure 13 Comparison of the MAPE between forward greedy and combinatorial search with a given number of additional features on top of the features, Re , c_{cf}^* , β , which form the prior knowledge.

VERIFICATION OF THE IMPORTANCE OF THE PRIOR KNOWLEDGE.

After identifying the additional important features on top of three features that had been chosen as prior knowledge, the assumption that the three prior knowledge features are actually important may be checked, by removing each of the them in turn from the top 5 feature subset (prior knowledge and A_{il}^*, κ_{il}).

The left side of Figure 14 shows the prediction error after all the 3 prior knowledge parameters, Re , c_{cf}^* and β are removed. The error is 26%. On the right is the prediction error with none removed (around 13%). The three data points in the middle of the figure show the effect of individually removing each feature one at a time. The three different lines compare the errors for the entire combined dataset with the subset of cases corresponding to the uniform and sheared flow tests.

When all the prior knowledge is removed, the average prediction error for cross-flow amplitude in both the uniform flow and sheared flow cases increases to about 25%.

When only the shear parameter β is removed, the averaged MAPE remains constant. This suggests that β was a poor choice to be included as prior knowledge variable. The small effect of β on mean error suggests that this information is redundant in this dataset. β could be left out of the subset with little effect. One reason that β has little effect is that the damping parameter c_{cf}^* for this dataset is also correlated with whether or not the flow is sheared or uniform, with uniform and sheared flows having substantially different dimensionless damping values, and hence provides the similar information to the prediction.

Finally, we observe that the Reynolds number, Re , is important in the uniform flow cases, but much less so in sheared flow.

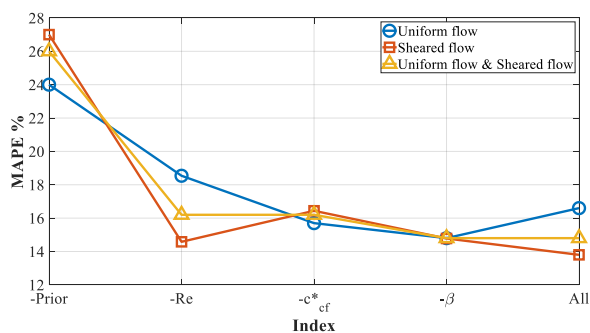


Figure 14 The comparison of MAPE before and after the prior knowledge of Re , c_{cf}^* or β are removed from the top 5-feature-subset

CONCLUSIONS

In this paper, following a formal definition of feature importance (Blum, 1997), a machine learning approach was used to identify important features for cross-flow VIV amplitude prediction.

The goal of this investigation was not just to develop a machine learning prediction algorithm, but rather to identify which parameters are necessary for robust predictions and hence identify which of the many parameters are most influential on the measured CF amplitude. Several conclusions may be drawn from the work:

1. Prior knowledge may be used to effectively constrain the machine learning model to aid in the search for additional important parameters.
2. The 'forward greedy' search approach to feature selection was shown to be much more efficient than brute force combinatorial searches.
3. The widely used damping parameter $m^*\zeta$ was shown to be not helpful for improving prediction accuracy.
4. The machine learning predictions showed strong correlation between in-line and cross-flow vibration amplitude for many test cases.
5. Reynolds number and the damping parameter c_{cf}^* were good choices as parameters to be selected as prior knowledge, whereas the shear parameter, β , was shown to be not important in the prediction of RMS cross-flow response for this dataset. This may not be true with other datasets that include many different types of current profiles.
6. Machine learning techniques are useful in determining the importance of various physical parameters in response prediction. This insight may prove useful in the challenge of using sparse measurements on full scale risers to determine fitness for service or attempt to use sparse data to justify the extension of projected fatigue life.

ACKNOWLEDGEMENT

This research has been sponsored by the members of the SHEAR7 Joint Industry Project: BP, Chevron, ExxonMobil, Petrobras, SBM Offshore, Shell, Equinor (Statoil) & Technip USA.

REFERENCES

- [1] Bourguet, R., Karniadakis, G., Triantafyllou, M., 2011. Vortex-induced vibrations of a long flexible cylinder in shear flow. *Journal of Fluid Mechanics* 677:342-382.
- [2] Blum, A.L. and Langley, P., 1997. Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1-2), pp.245-271.
- [3] Bishop, Christopher M. *Pattern recognition and machine learning*. springer, 2006.
- [4] Dahl, J. J. M. 2008. Vortex-induced vibration of a circular cylinder with combined in-line and cross-flow motion (Doctoral dissertation, Massachusetts Institute of Technology).
- [5] Govardhan, R. N., & Williamson, C. H. K. (2006). Defining the 'modified Griffin plot' in vortex-induced vibration: revealing the effect of Reynolds number using controlled damping. *Journal of Fluid Mechanics*, 561, 147-180.

- [6] Griffin, O.M., Skop, R.A. and Ramberg, S.E., 1975, January. The resonant, vortex-excited vibrations of structures and cable systems. In Offshore Technology Conference. Offshore Technology Conference.
- [7] Guyon, I. and Elisseeff, A., 2003. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), pp.1157-1182.
- [8] Lie, H. et.al., 2013, August. Comprehensive riser VIV model tests in uniform and sheared flow. In ASME 2012 31st International Conference on Ocean, Offshore and Arctic Engineering. 923-930.
- [9] Marcollo H., Resvanis T.L., Dillon C., Kilner A., Vandiver J.K., SHEAR7 V4.10 Userguide, 2018.
- [10] Modarres-Sadeghi, Y., Chasparis, F., Triantafyllou, M. S., Tognarelli, M., & Beynet, P. (2011). Chaotic response is a generic feature of vortex-induced vibrations of flexible risers. *Journal of Sound and Vibration*, 330(11), 2565-2579.
- [11] Rao, Z., Resvanis, T.L. and Vandiver, J.K., 2014, June. The identification of power-in region in vortex-induced vibration of flexible cylinders. In ASME 2014 33rd International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers Digital Collection.
- [12] Rao, Z., 2015. The flow of power in the vortex-induced vibration of flexible cylinders. Ph. D. Dissertation, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- [13] Resvanis, T.L. et.al., 2012, July. Reynolds number effects on the vortex-induced vibration of flexible marine risers. In ASME 2012 31st International Conference on Ocean, Offshore and Arctic Engineering. 751-760.
- [14] Sarpkaya, T. 2004. A critical review of the intrinsic nature of vortex-induced vibrations. *Journal of fluids and structures*, 19(4): 389-447.
- [15] Scruton, C. 1955. Wind-excited oscillations of tall stacks, *Engineer* 199. 806-808.
- [16] Swithenbank, S.B., Vandiver, J.K., Larsen, C.M. and Lie, H., 2008, January. Reynolds number dependence of flexible cylinder VIV response data. In ASME 2008 27th International Conference on Offshore Mechanics and Arctic Engineering (pp. 503-511). American Society of Mechanical Engineers.
- [17] Vandiver, J. K. 1993. Dimensionless parameters important to the prediction of vortex-induced vibration of long, flexible cylinders in ocean currents. *Journal of Fluids and Structures*, 7(5), 423-455.
- [18] Vandiver, J.K., 2012. Damping parameters for flow-induced vibration. *Journal of fluids and structures*, 35, pp.105-119.
- [19] Vandiver, J.K. and Ma, L., 2017, June. Does More Tension Reduce VIV?. In ASME 2017 36th International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers Digital Collection.
- [20] Vandiver, J.K., Ma, L. and Rao, Z., 2018. Revealing the effects of damping on the flow-induced vibration of flexible cylinders. *Journal of Sound and Vibration*, 433, pp.29-54.

Table 1 Test matrix for 38-m-long flexible cylinder's VIV experiment

Pipe	No. of flow speeds	Current profile	Re	Pretension P_0 (N)	Vr_{cf}	Dominant cross-flow mode
30 mm Pipe	12	Uniform	$1.5 \times 10^4 - 5 \times 10^4$	1500-2000	6.0-8.0	5-11
30 mm Pipe	18	Linear shear	$1.5 \times 10^4 - 7.5 \times 10^4$	1800-2000	5.6-8.9	5-13
80 mm Pipe	10	Uniform	$7.2 \times 10^4 - 1.5 \times 10^5$	2000-6000	6.5-8.0	4-12
80 mm Pipe	12	Linear shear	$6 \times 10^4 - 1.5 \times 10^5$	2000-6000	6.2-9.0	4-8

Table 2 List of top parameter combinations that give the smallest prediction errors with 1-4 additional parameters

Number of added features m	Added Feature on top of the prior knowledge	Δ MAPE	MAPE(m)	MAPE for the sheared flow cases	MAPE for the uniform flow cases
1	A_{il}^*	6.0%	18.0%	14.5%	23.4%
1	κ_{il}	3.4%	20.6%	18.8%	23.1%
2	A_{il}^*, κ_{il}	3.2%	14.8%	13.8%	16.6%
2	A_{il}^*, κ_{cf}	2.7%	15.3%	13.4%	18.1%
3	$A_{il}^*, \kappa_{il}, \alpha_{cf}$	0.8%	14.2%	13.7%	14.7%
3	$A_{il}^*, \kappa_{il}, Vr_{il}$	0.7%	14.3%	13.3%	15.6%
3	$A_{il}^*, \kappa_{il}, Lk_{cf}$	0.7%	14.3%	13.9%	15.0%
4	$A_{il}^*, \kappa_{il}, Lk_{cf}, \alpha_{cf}$	1.1%	13.0%	12.4%	13.9%
4	$A_{il}^*, \kappa_{il}, Lk_{cf}, Vr_{il}$	1.1%	13.0%	12.5%	13.9%
4	$A_{il}^*, \kappa_{il}, \kappa_{cf}, Vr_{il}$	1.0%	13.1%	11.8%	14.9%