Improved Damage Detection in Pelton Turbines Using Optimized Condition Indicators and Data-driven Techniques

Weiqiang Zhao[1], Mònica Egusquiza[1], Aida Estevez[2], Alexandre Presas[1], Carme Valero[1], David Valentín[1], Eduard Egusquiza[1]

[1] *Centre for Industrial Diagnostics and Fluid Dynamics (CDIF), Polytechnic University of Catalonia (UPC), Barcelona, Spain, weiqiang.zhao@upc.edu, +34 657080743*
[2] *Dpto. Ingeniería mecánica y fabricación. Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Sevilla, Spain*

Abstract

The health condition of hydraulic turbines is one of the most critical factors for the operation safety and financial benefits of a hydro power plant. After the massive entrance of intermittent renewable energies, hydropower units have to regulate their output much more frequently for the balancing of the power grid. Under these conditions, the components of the machine have to withstand harsher excitation forces, which are more likely to produce damage and eventual failure in the turbines. To ensure the reliability of these machines, improved condition monitoring techniques are increasingly demanded.

In this paper, the feasibility of upgrading condition monitoring of Pelton turbines using novel vibration indicators and data-driven techniques is discussed. The new indicators are selected after performing a detailed analysis of the dynamic behavior of the turbine using numerical models and field measurements. After that, Factor Analysis (FA) is carried out in order to assess which are the most informative indicators and to reduce the dimension of the input data.

For the validation of the proposed method, monitoring data from an actual Pelton turbine that suffered from an important fatigue failure due to a crack propagation on the buckets have been used. The novel condition indicators as well as classical indicators based on the spectrum and harmonics levels have been obtained while the machine was in good operation, during different stages of damage and after repair. All of these have been used to train an Artificial Neural Network (ANN) model in order to predict the evolution of the crack until failure occurs. The results show that using the improved monitoring methodology enhances the ability to predict the appearance of damage in comparison to typical condition indicators.

Keywords: condition monitoring; Pelton turbine; damage detection; condition indicator; Factor Analysis (FA); Principal Component Analysis (PCA).

1. Introduction

Hydropower is one of the most important renewable energies, which converts the potential energy of water into electric energy. After its development since more than one century ago, hydropower has become a reliable energy resource. Over the last years, with the entrance of new renewable energies like solar photovoltaics and wind energy, whose output is random and hard to regulate, hydropower has become extremely significant for balancing the generation and consumption of electricity. Over the last decade, the installed capacity of wind and solar energy has increased from 13.2% and 2.1% to 24.0% and 20.7%, respectively [1]. This new scenario requires hydropower units to regulate their output much more frequently for the balancing of the power grid. Therefore, more flexibility is demanded and turbines have to operate many hours off-design. Due to the more severe operating conditions, components from Pelton turbine are more prone to suffer damage and/or failure [2–4]. Unexpected failures of the turbine may result in substantial economic losses, not only due to the costs related to the repair or replacement of the component but also to the downtime and production loss thus caused.

As the most common action-type turbine, Pelton turbines are widely used in high-head hydropower plants. In the turbine, the high pressure of the water at the entrance is converted into a high velocity jet using a nozzle. The jet impinges the runner as shown in Figure 1, converting the kinetic energy of the water into mechanical energy that is converted into electricity by the generator. During operation, strong pulsating forces are produced on the runner. The buckets that receive directly the high speed jet of water resemble a cantilever beam and have to transmit the torque to the wheel. For this reason, the fatigue of the material on the bucket area is one of the most common causes of failure in this type of turbines.

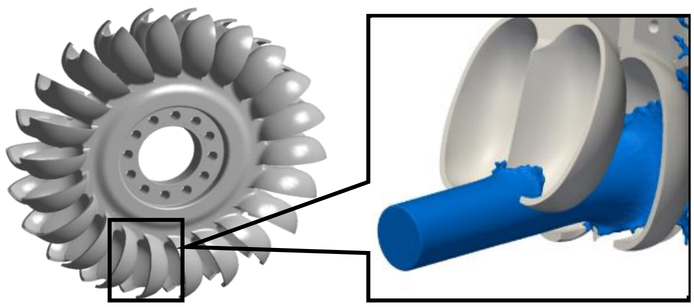


Figure 1 Schematic of the operation of the Pelton turbine bucket[5]

In order to ensure the availability of the machine, effective strategies such as periodic inspections and condition monitoring have been applied for protecting the machine and preventing serious failure [6]. Usually, sensors are installed on bearings where the vibrations generated by the turbine and the generator are transmitted. Vibration signals, as well as the operating parameters, are recorded by an acquisition system to be analyzed for diagnosis. From vibration signals condition indicators can be calculated, which should be effective to detect damage. These indicators are trended and compared to some alarm and trip levels.

Nowadays, remote online monitoring systems have been applied on hydropower plants so that it is possible to monitor the turbine in a diagnostics center far away from the power plant [7]. Accelerometers located in bearings are the sensors currently used to measure machine vibrations. Other types of sensors like acoustic emission sensors, strain gauges and microphones have also been tested to monitor the condition of hydraulic turbines[8–10].

Selecting alarm and trip levels able to identify incipient damage in the runner is a complex task[6]. First, the levels of each indicator must be mapped for all operating conditions with the machine in good condition. Second, a mapping of the evolution of these condition indicators under a damage situation should also be represented. Machine learning techniques can be used for such purpose[11], but they need large amounts of historic data that encompass the evolution of the machine vibration from good condition to failure. Since turbines are never allowed to operate until the end of their useful life, one of the limitations in condition monitoring of hydropower plants is the lack of data. These shortcomings can be partially overcome with the use of sophisticated numerical simulation models where synthetic damage can be simulated[12].

Being different from the general machinery, the main components of hydro turbine units such as the runner and the generator are tailor-made pieces. The layout of turbine units also varies from one machine to another, which leads to different structural responses. Because of the above reasons, the diagnosis of a hydraulic turbine depends much on expertise and historical data. Some typical types of damage can be detected but many of them, especially the ones taking place in the runner, are not detectable before failure occurs[13]. This happens because the runner vibrations can hardly be detected by monitoring vibrations in the bearing; natural frequencies of the runner do not produce important deformations in the rotor[14].

In recent years, significant efforts have been made in application of data-driven and artificial intelligence (AI) for structural damage detection. Several researchers have applied Artificial Neural Network (ANN) on the vibration pattern for damage diagnosis or fault detection on various types of machines including helicopter pump [15], wind turbine [16], railway wheel [17], etc. However, there are a few studies on the implementation of AI on the experimental data of hydraulic turbines. With the development of computer techniques, more simulation data on hydraulic turbines are available for the training of ANN: R.A. Saeed et al. [18] used Principal Component Analysis (PCA) to extract features from the frequency response function (FRF) obtained from the simulation of a hydro turbine runner. The FRF features were used for training ANN in order to predict the crack length on the blade. With the years of accumulation of on-line condition monitoring data, it would become possible the application and optimization of data-driven and AI techniques for the monitoring of hydraulic turbines.

Although AI has been proven to be an effective tool for vibration-based structural damage diagnosis in several studies, there remains a very common hurdle in most of the studies: the size of the input data, which is determined by the resolution of the spectrum. A higher resolution provides more details about the dynamic response of the structure but the amount of data is too large for the neural networks' applications on engineering problems. A general way to deal with this problem is to use principal component analysis (PCA) to reduce the size of the input pattern[13]. However, the calculation of one principal component (PC) depends on every input variable, which means the PCs will change with the increase of the samples. In addition, there is no real meaning in the extracted principal components [14]. To address this problem, factor analysis is proposed. It describes variables in terms of a lower number of factors that have different loading values regarding the input variables and also tries to find out one or more latent variables (factors) that exert causal influence (loadings) on these observed variables [15]. It was firstly proposed in 1904 for the explanation of psychological theories [16]. During its development of more than one century, many researchers have contributed to its theory including the number of factors to retain [17], factor extraction algorithms [18] and so on. Nowadays, factor analysis has been applied into engineering applications: Mansi Tripathi and Sunil Kumar Singal introduced factor analysis into weight determination to develop a novel water quality analysis[19]. Reenu Maskey et al. provided explicit information on the appropriate application of exploratory factor analysis on engineering problem[20]. The development of factor analysis has made possible to optimize monitoring by compressing the monitoring indicators into interpretable factors.

In this paper, a novel condition monitoring method for Pelton turbines is introduced and discussed. The first step consists in defining optimized condition monitoring indicators based on an extensive analysis of the dynamic behavior of a Pelton turbine. The main advantage of these new condition indicators is the ability of detecting runner vibrations and its variation with abnormal operation and damage. The second step consists in applying FA in order to optimize the new indicators and reduce their dimension.

To validate this new procedure, monitoring data from an existing Pelton turbine has been used. The machine suffered a catastrophic failure in which one of the buckets broke off during operation. The inspection of the machine later revealed the failure was due to a deviated jet, which caused several cracks to appear and propagate on different buckets. To check the ability of the new methodology to predict the propagation of the cracks, the novel condition indicators have been extracted from the available data and used to define an Artificial Damage Index (*ADI*) based on the typical crack propagation behavior. These have been used in an ANN model in order to predict the *ADI* until failure occurs. The same process has been followed for the typical condition indicators used normally in condition monitoring of Pelton turbines in order to compare the results with the new method.

The paper is organized as follows: Section 2 introduces the novel condition indicators and the FA algorithm, Section 3 presents the application of the proposed method and two comparative methods on the real case, in Section 4 the failure prediction performance of the proposed indicators is compared with the conventional methods and Section 5 provides the concluding remarks.

2. Methodology for improved monitoring of Pelton turbines

2.1. Definition of new condition indicators based on dynamic analysis

Typical vibration condition indicators are effective for detecting some types of damage, but in many cases the symptoms are only perceptible when damage is at an advanced stage [2,29,30]. It is thus of high interest to upgrade the current monitoring procedures so that any variation in the behavior of the turbine is rapidly detected by the system and the potential effects on the lifetime of the structure assessed. In Figure 2, a sketch of a typical Pelton turbine has been represented. The excitation force is applied to the runner and vibrations are measured in the bearings with accelerometers.

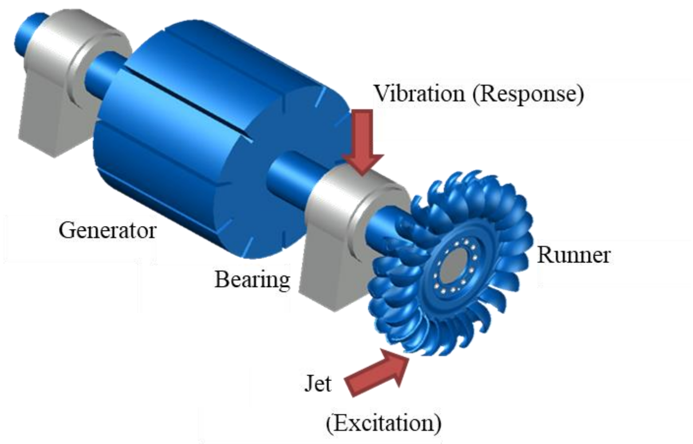


Figure 2 Pelton sketch showing the excitation and the monitoring position

As indicated above, the jet impinges directly on the runner buckets generating strong vibrations that depend on the turbine structural response. In this type of turbine, the natural frequencies more prone to be excited are the ones of the runner. One of the issues, as in many hydraulic turbines, is that runner vibrations can hardly be detected by monitoring vibrations in the bearing because natural frequencies of the runner do not produce important deformations on the rotor. In Figure 3, the deformations in a runner and rotor when excited by the jet have been represented. The numerical simulation was done with Finite Element Method (FEM) and checked experimentally. It can be observed that while the runner has large deformations, the rotor is barely affected. In case of damage or abnormal operation, the runner dynamics varies and these changes have to be detected by the monitoring system.

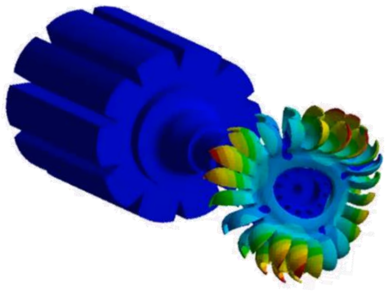


Figure 3 Deformations of a Pelton runner

A more detailed dynamic analysis of Pelton runners [31] indicate that the runner response is quite complicated, with many natural frequencies and mode-shapes (tangential, axial and radial). They are close to each other with high modal density and covering a high frequency range well above the rotor natural frequencies. In Figure 4, the first natural frequencies of a Pelton runner and associated mode-shapes detected during the start-up transient are indicated.

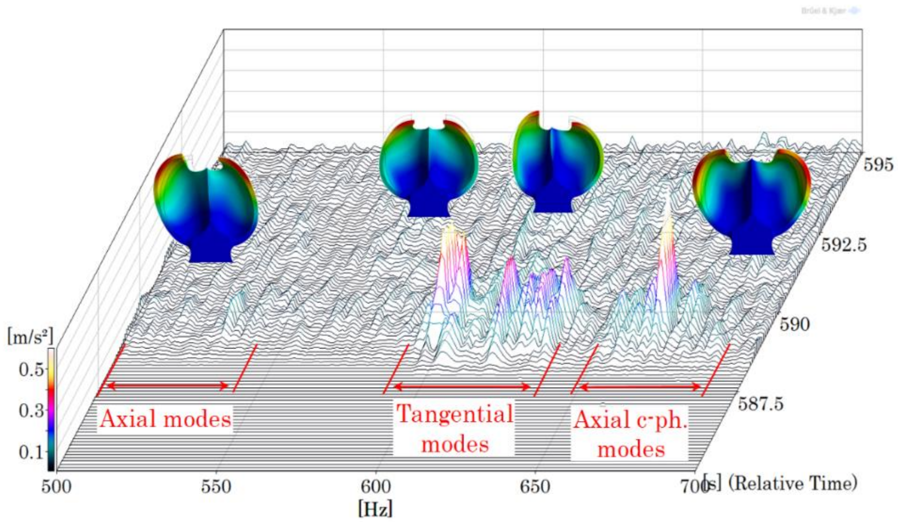


Figure 4 Excitation of the first runner natural frequencies during start-up of a Pelton turbine

The main forces affecting a Pelton turbine during operation are of mechanical, hydraulic and electromagnetic origin. The main hydraulic force comes from the impingement of the water jet on the buckets. Every time a bucket is in front of an injector, the runner receives a strong impact. During the impact, runner natural frequencies are excited. In normal operating conditions the tangential modes are basically the ones to be excited (jets hit the runner tangentially), but with abnormal operation or incipient damage other types of natural frequencies are excited too. The emergence of incipient damage in a structure may lead to a change in its dynamic response, which subsequently can be reflected in the vibration behavior [13,32]. The point is to detect the change in the runner vibrations and associate them to abnormal operation and damage.

For improving the condition monitoring the natural frequencies of the rotor and of the runner have to be monitored. In the new condition indicators, besides the typical synchronous bands for unbalance, misalignment and bucket passing frequencies, other bands related to the natural frequencies of the structure have been considered.

2.2. Dimension reduction of indicators based on Factor Analysis and Principal Component Analysis

In this paper two different methods are used to reduce the dimension of the indicators sets. These methods are the so-called Principal Component Analysis (PCA) and Factor Analysis (FA).

The procedure of PCA is shown in Figure 5. By PCA, an indicator matrix is rotated into a new projection matrix. The projection vectors (each being a linear combination of the variables) are an uncorrelated orthogonal basis set. The first PC has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. Only the components with highest eigenvalues are retained for the analysis[22].

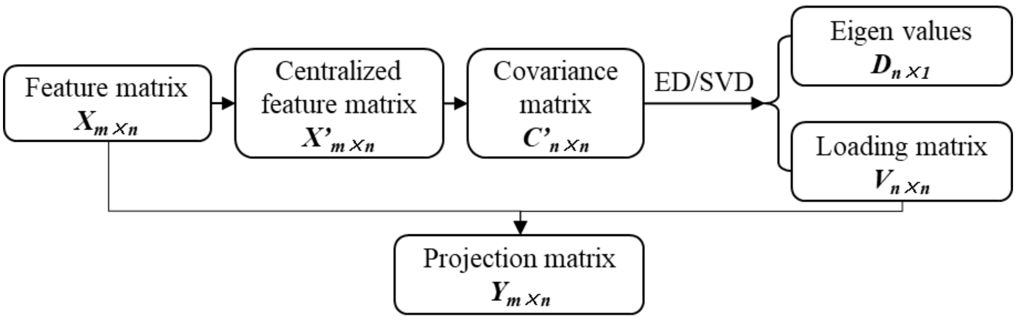


Figure 5 Procedure of PCA

In a FA model, each condition indicator can be represented as a linear combination of the common factors and specific factors[33]:

$$X_{d\times n} = \mu_{d\times n} + \Lambda_{d\times M} F_{M\times n} + E_{d\times n} \tag{1}$$

Where X is the indicator matrix with n indicators (n dimension) and d samples, μ is the mean value of each column of X and Λ is a constant d-by-M matrix of factor loadings. The $(i,j)^{th}$ element of the d-by-M matrix Λ is the coefficient, or loading, of the j^{th} factor for the i^{th} variable. F is a matrix of M common factors of the indicator matrix, and E is a vector of independent specific factors which represents the portion that cannot be explained by the common factors. In the indicator matrix, all of the indicators share these M common factors and each can be represented as a linear function of common factors and its specific factor:

$$X_i = \mu_i + a_{i1}F_1 + a_{i2}F_2 + \cdots + a_{iM}F_M + \varepsilon_i, (i = 1,2, \ldots d) \quad (2)$$

The indicators are related to the common factors by the factor loadings and can be classified into different common factors according to the loadings. Moreover, each common factor usually has a physical explanation regarding the dynamic behavior of the machine, which is an important advance compared to PCA. The main steps of FA are the following:

1) Calculating the eigenvalue of the covariance matrix of the input samples. The eigenvalues are the key for deciding the number of the common factors. According to the literature reviewed, a scree test is the most common and accurate criterion [34]: retaining all the factors above (i.e., to the left of) the inflection point of the scree plot of the eigenvalues.

2) Loading matrix calculation. With damage, some condition indicators severely change their values, especially when the machine approaches failure. Therefore, principal factor analysis is applied in the proposed method, which is suitable for corrupt data[23]. According to principal component method, the loading matrix Λ is computed by the following equation:

$$\Lambda_{d\times M} = \left(\sqrt{\lambda_1}\eta_1, \sqrt{\lambda_2}\eta_2, \cdots, \sqrt{\lambda_M}\eta_M\right) \quad (3)$$

Where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$ are the m largest eigenvalues of the correlation coefficient matrix of X and η_1, η_2, …, η_M are the corresponding orthonormalized eigenvectors of the eigenvalues.

3) Factor rotation. The idea of factor rotation is to transform the loading matrix into a structure where each of the retained factors are ideally loaded on fewer variables, i.e. to make sure the projection of each variable on the rotated factor axis is the largest (or the lowest). By factor rotation, the new factor loading matrix can be simplified and the factors have more physical meaning or interpretability.

2.3. Failure prediction by ANN regression

As one of the most common regression tools, a supervised ANN is developed in order to build up the relation between condition indicators and damage. Since a single hidden layer ANN with sufficient number of neurons will be suitable to fit any continuous function[35], a feed-forward neural network with one hidden layer as shown in Figure 6 has been used. The input layer consists of neurons for the normalized monitoring indicators, and the neurons in the hidden layer are defined as sigmoid activation functions. The output layer contains one neuron which equals to a parameter related to a specific damage such as crack length.

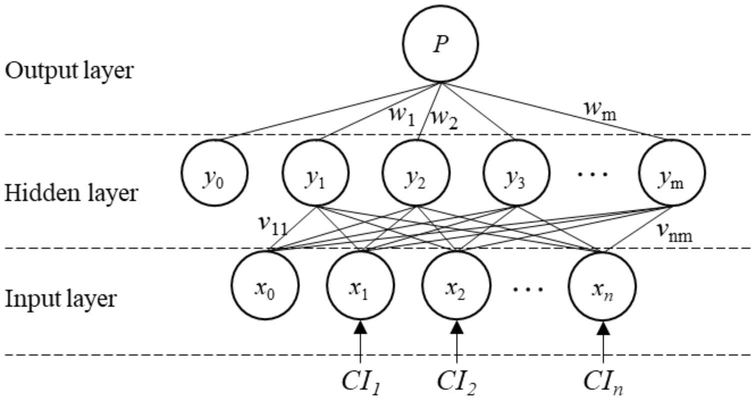


Figure 6 Architecture of the ANN model

The LM (Levenberg-Marquardt) backpropagation function is applied as the update function:

$$\Delta w = [J^T J + \lambda I]^{-1} J^T e \tag{4}$$

Where J is the Jacobian matrix of the partial derivative of E with respect to v or w. I and e represent an identity matrix and output error vector respectively. λ is large in the beginning of the learning phase, which increases the convergence speed and then decreases in order to make the approximation more accurate. The ANN model will be trained by the indicators and be used for damage detection.

3. Study on an actual Pelton turbine

3.1. Case description

An existing Pelton turbine prototype has been investigated to verify the proposed method. It is a horizontal shaft machine consisting of two runners, one shaft and a generator. The structure is supported by two bearings, located between each one of the runners and the generator. Both runners have 22 buckets and each one of them is operated by one jet. The main characteristics of the machine are listed in Table 1.

Table 1 Basic operation parameters of the turbine unit

Parameter	Value
Rated head	770 m
Maximum power	34 MW
Rotating speed f_f	600 rpm (10Hz)
Number of buckets Z_b	22
Bucket passing frequency f_b	220 Hz

The researched turbine had been monitored by a monitoring system for 13 years. A total of 5 accelerometers were installed on both bearings in the radial and axial directions, as seen in Figure 7.

Accelerometers A11 and A31 were installed horizontally in the same direction in which the jet impinges the runner. A14, A34 and A35 were installed in the vertical and axial directions, respectively. The data was collected when the machine was operating at 40% and 100% of the rated output. Therefore, the measurement points are named after sensor and output (e.g. A14-04 and A14-10 represent measurement point A14 at 40% and 100% rated output respectively).

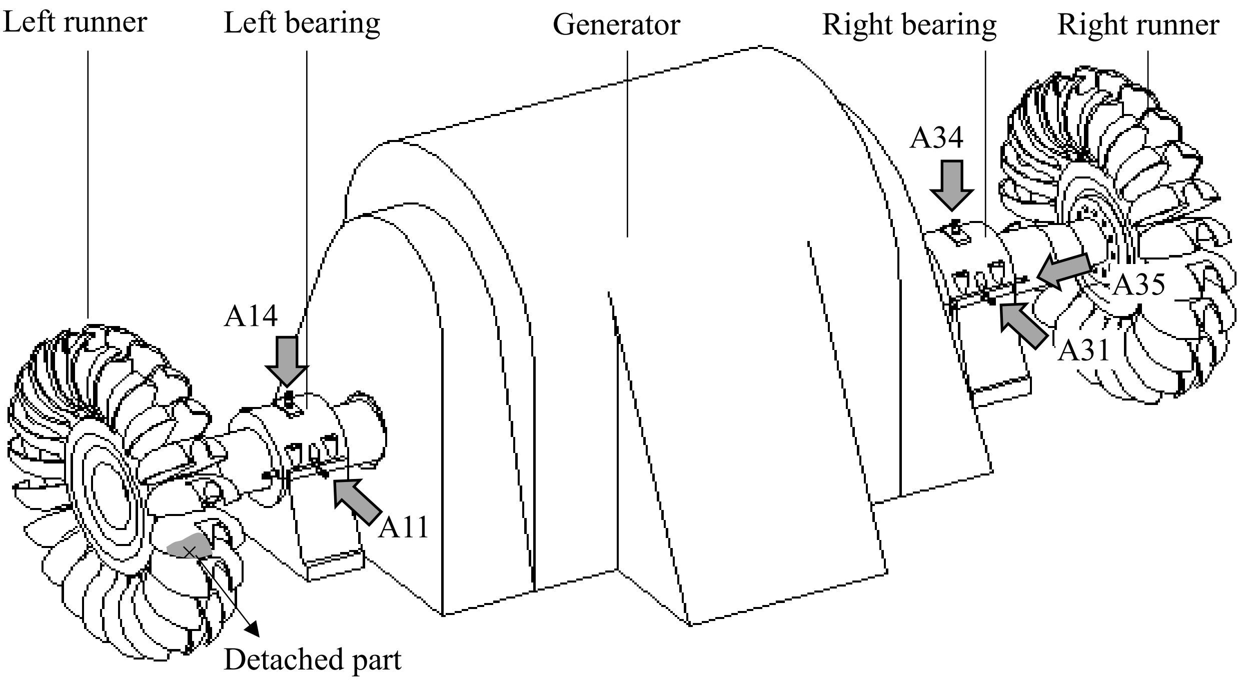


Figure 7 Sketch of the Pelton turbine and layout of the vibration sensors

For the analysis, a history case has been selected. On the 29th of August, 2011, the system detected an abnormal root mean square (RMS) vibration value surpassing the alarm threshold and the machine was stopped. The following inspection revealed the cause of the increased vibration: a piece of one bucket broke off while in operation and, in addition, cracks were present in the same area of several buckets (Figure 8). The inspection of the runner indicated that the incident took place due to the fatigue of the material, for which it was assumed that the machine had not been operating properly for some time. Machine vibrations in all the monitoring locations were available with the machine in good condition, with incipient damage and with severe damage [36,37]. After that, the machine was repaired and put again into service.

Figure 8 Damage in a bucket

3.2. New condition monitoring method

In order to implement the new condition indicators on the monitoring system, an extensive analysis of the dynamic behavior of the turbine was carried out [31,39]. First, the natural frequencies of the machine were determined by means of numerical and experimental analysis on the prototype. After that, the response (i.e., vibration) of the machine under the dynamic loading of the water jets was studied. These investigations provided a deeper understanding of the operation of the Pelton turbine and allowed defining new frequency bands related to the most excitable modes of the runner. Besides the typical synchronous bands for unbalance, misalignment and bucket passing frequencies, other bands related to the natural frequencies of the structure were devised. The latter is divided into two main groups, the ones corresponding to the rotor's natural frequencies and the ones belonging to the runner.

The rotor modes involve the deformation of the whole turbine and are found at frequencies below 300 Hz. The runner modes are found over 500 Hz and only involve the deformation of the runner and/or the buckets. Only the most relevant harmonic bands are selected according to the main excitation forces on the Pelton turbine (f_f,$2f_f$, $3f_f$, f_b, $2f_b$ and $3f_b$) (Figure 9). The proposed new condition indicators are listed in Table 2. For each measuring position, 12 spectral bands are selected as damage indicators.

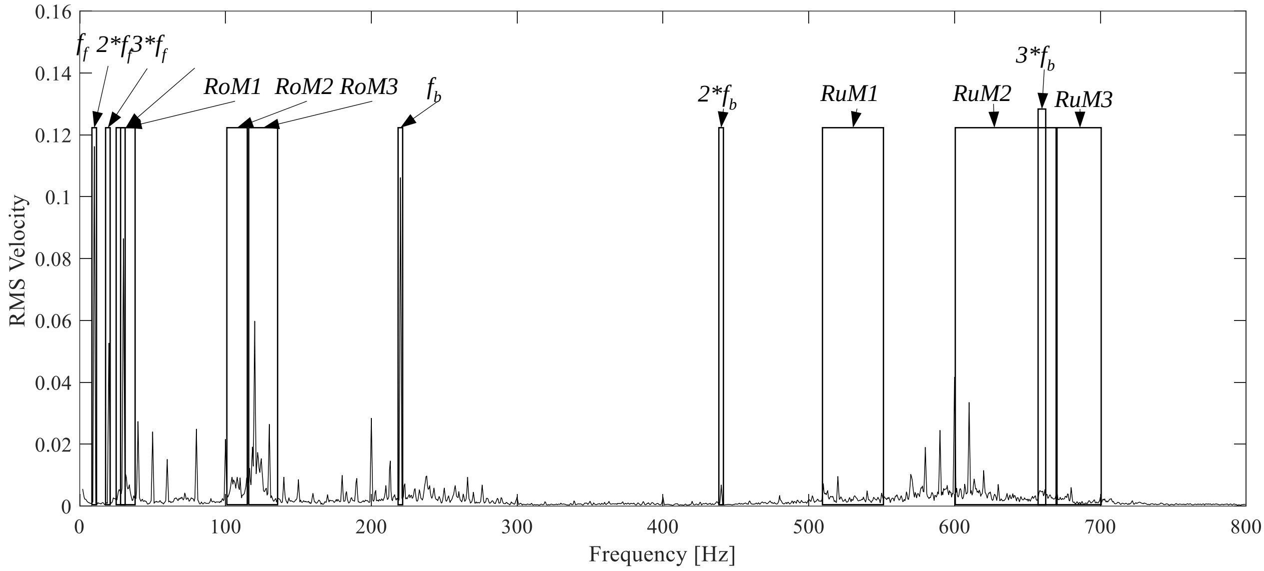


Figure 9 Vibration spectrum of the Pelton turbine and spectral bands

Table 2 The frequency ranges of the spectral bands

Mode/ frequency	Mode description	Abbreviation	Frequency range
	f_f	f_f	9-11Hz
f_f and harmonics	2*f_f	2f_f	19-21Hz
	3*f_f	3f_f	29-31Hz
	Rotor mode 1	*RoM*1	25-38Hz
Rotor modes	Rotor mode 2	*RoM*2	102-115Hz
	Rotor mode 3	*RoM*3	115-134Hz
	f_b	f_b	219-221Hz
f_b and harmonics	2*f_b	2f_b	439-441Hz
	3*f_b	3f_b	659-661Hz
	Axial runner mode	*RuM*1	510-550Hz
Runner modes	Tangential runner mode	*RuM*2	600-670Hz
	Rim runner mode	*RuM*3	670-700Hz

After extracting the new condition indicators based on the dynamics of the turbine, FA technique is used to reduce the dimension of the database and select the most relevant indicators. FA is preferred over PCA for dimension reduction as the physical meaning of the indicators is retained. The measurement point A34-10 is taken as an example to describe the process.

First, the number of common factors is determined by scree test[26]. The relative eigenvalues, as well as their cumulative percentages, are displayed in Figure 10. From the scree plot it can be seen that the inflection point is the third one, which indicates that three common factors are sufficient for retaining most of the information of the indicator matrix.

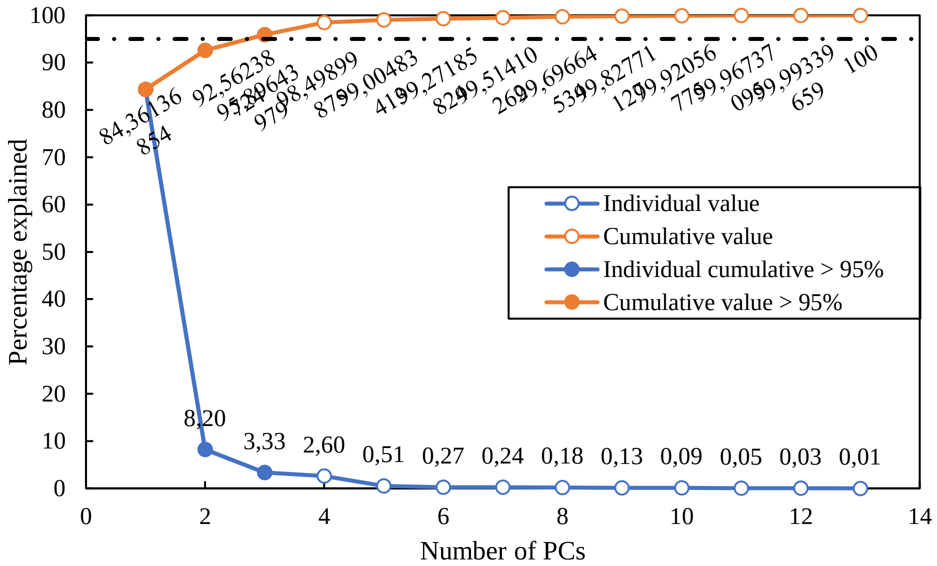


Figure 10 Scree plot of the PCA result of point A34-10

After the retained number has been determined, FA is conducted on the indicator matrix. The loading matrix is calculated according to Equation 3. The loadings of each indicator on the 3 common factors are listed in Table 3, as well as the module of each indicator, which is the Euclidean norm of the loadings.

Table 3 Loading factor of each damage indicator of point A34-10

Indicator	Factor1	Factor2	Factor3	Module
f_r	-0.4337	0.1769	0.3147	0.5643
$2*f_r$	0.0387	-0.0440	-0.8386	0.8406
$3*f_r$	-0.3144	0.8485	0.3779	0.9806
Rotor mode 1	-0.2751	0.9032	0.3218	0.9975
Rotor mode 2	0.0368	0.4883	0.0200	0.4901
Rotor mode 3	-0.4351	0.2472	0.3843	0.6310
f_b	-0.0392	0.1362	0.2187	0.2606
$2*f_b$	-0.4199	0.5066	0.7337	0.9855
$3*f_b$	0.9085	0.0224	-0.0414	0.9097
Runner mode 1	-0.4024	0.6606	0.5677	0.9595
Runner mode 2	0.8938	-0.1856	-0.2492	0.9463
Runner mode 3	0.8416	-0.1578	-0.0389	0.8571

Taking advantage of 3 loading factors, the loading matrix of point A34-10 can be represented in a 3D plot, which is helpful for visualizing and comparing the loadings of the indicators. The loading vectors are shown in Figure 11. Thus, the direction and length (module) of each vector indicate its relevancy with the common factors. For example, there are 4 indicators with negative loadings on the first common factor and 7 with positive loadings. The indicator $3*f_b$ has the smallest angle with the axis and largest loading value, which indicates $3*f_b$ has the greatest dependency with the first common factor. On the contrary, the length of vector

f_b is 0.2606, which is too small to represent any common factor.

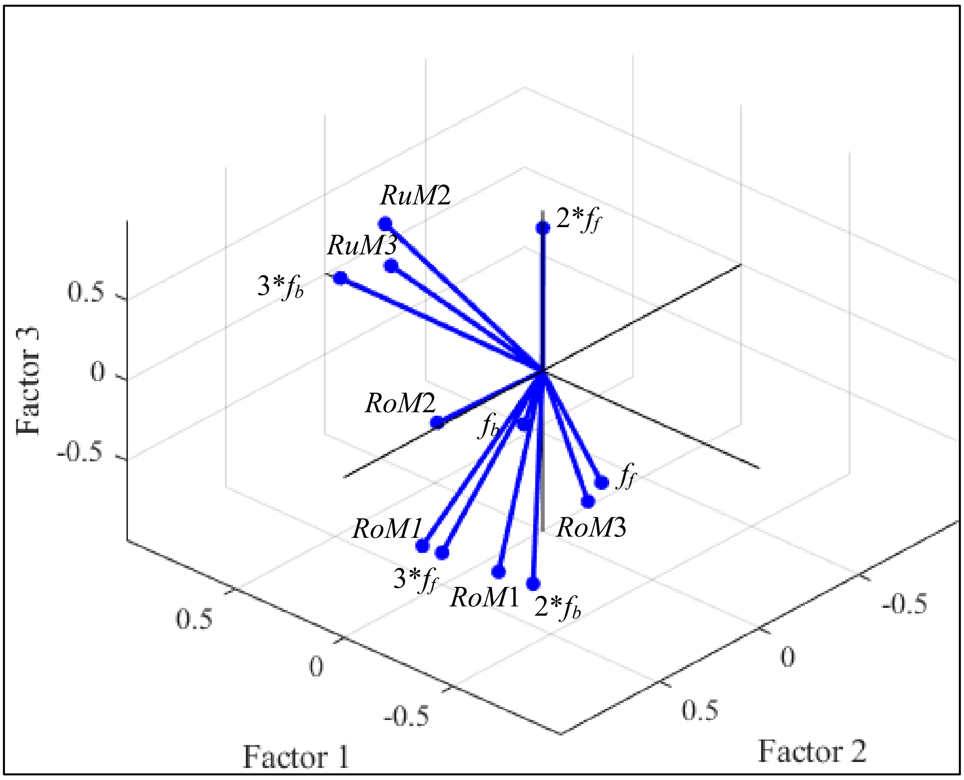
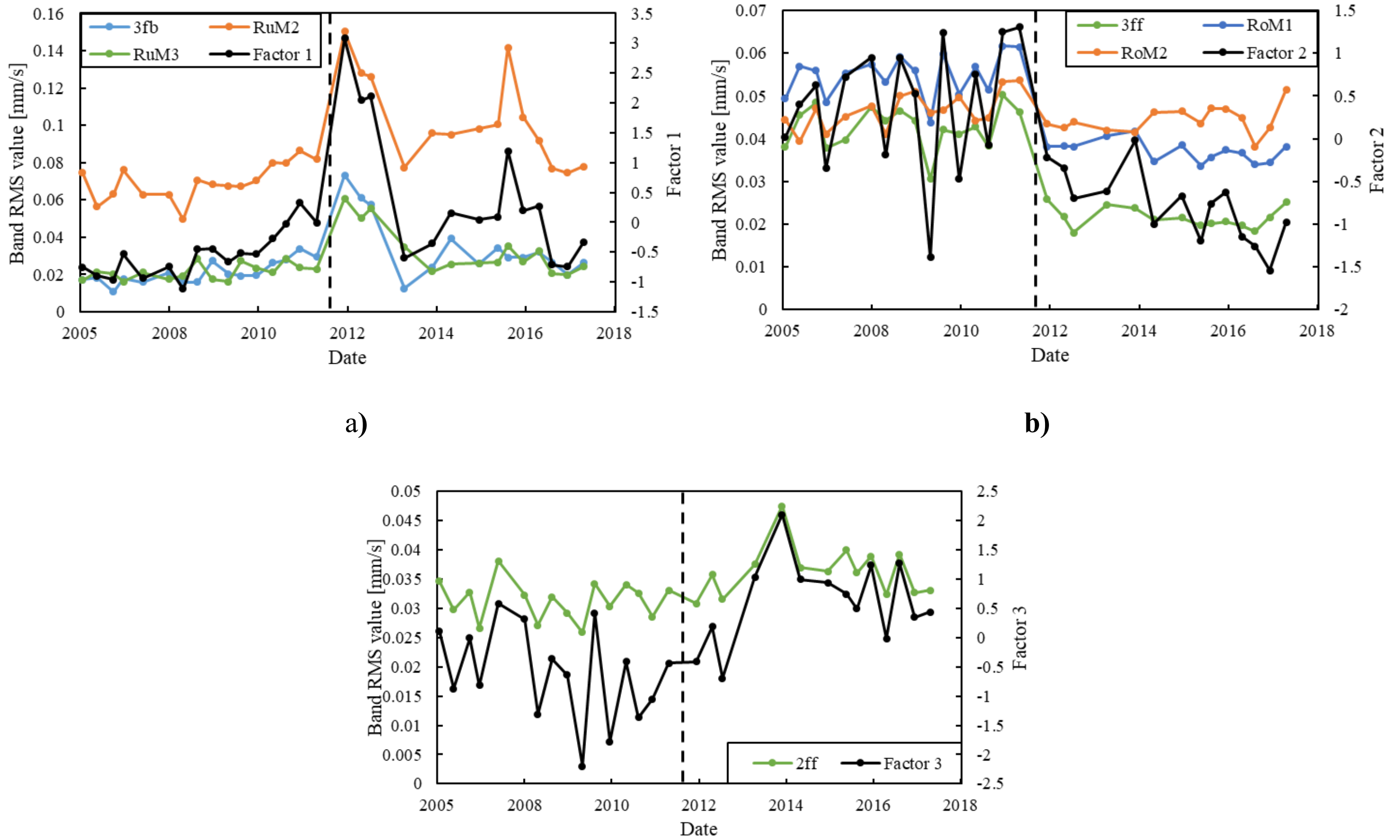


Figure 11 Loading factors of each new indicator on the three common factors

The three common factors are calculated and displayed in Figure 12, as well as their corresponding indicators. The data before and after repair is separated by a dash line. According to Table 3 and Figure 11, the indicators 3f_b and *RuM*2 and *RuM*3 have the highest loading value on the common factor 1, which means they have a similar trend with each other before and after failure. The evolution of these three indicators as well as the corresponding factor are shown in Figure 11 a. Before the failure their values grow gradually and after repair they increase dramatically. In the whole range, their values are higher after repair than before failure. However, factor 2 (Figure 11 b) shows an inverse trend to factor 1 since the value plummets after repair. Being Similarly, it shows a slight increase before failure, as factor 1. The indicators 3f_f, *RoM*1 and *RoM*2 have a high relevancy with this common factor. As for factor 3 (Figure 11 c), the value is higher in good condition than damage condition. However, there's no significant trend during failure.

The FA can be regarded as a modal decomposition process: the vibration behavior of the structure is a superposition of different modes. Therefore, the indicators with high loading on the common factors are the most significant indicators and the indicators which have low loading on every factor are regarded as redundant components. For the measurement points, the indicators with loading module larger than 0.5 are selected as the most significant indicators. The same process has been carried out on the indicators of every measurement point and the modules have been listed in Table 4. The indicators with modules lower than 0.5 (marked in red) were eliminated from the indicator matrix. In this way, the dimension of the indicator set is reduced.



a) b) c)

Figure 12 Common factors and the corresponding damage indicators

Table 4 Loading factor modules

Indicator	A11-04	A11-10	A14-04	A14-10	A31-04	A31-10	A34-04	A34-10	A35-04	A35-10
f_f	0.89	0.92	**0.45**	**0.37**	0.65	0.75	0.61	0.56	0.79	0.87
$2f_f$	**0.21**	0.91	0.53	0.86	0.86	0.93	0.62	0.84	0.53	0.87
$3f_f$	0.76	0.99	1.00	1.00	1.00	1.00	0.99	0.98	0.97	1.00
$RoM1$	0.77	0.96	0.91	0.89	1.00	1.00	1.00	1.00	1.00	0.99
$RoM2$	0.96	0.92	0.92	0.89	0.74	0.51	0.74	**0.49**	0.80	**0.22**
$RoM3$	0.78	0.93	0.91	0.95	0.95	**0.22**	0.92	0.63	0.94	0.54
f_b	0.69	0.69	0.75	0.87	0.90	0.77	0.61	**0.26**	**0.43**	**0.44**
$2f_b$	0.59	0.51	**0.48**	0.84	0.75	0.58	0.82	0.99	**0.33**	0.80
$3f_b$	**0.42**	**0.23**	**0.09**	0.54	0.97	0.61	1.00	0.91	0.88	0.84
$RuM1$	0.97	0.96	0.94	0.98	0.86	0.94	0.91	0.96	0.91	0.93
$RuM2$	0.99	0.93	0.80	0.89	0.58	1.00	0.77	0.95	0.94	0.97
$RuM3$	0.70	**0.44**	0.67	0.52	0.98	0.92	0.97	0.86	1.00	0.95

3.3. Comparative condition indicators (conventional indicators)

Two comparative sets of indicators are extracted from the raw signal and used for training the ANNs to compare their respective damage detection performance with the newly proposed method. In these two cases, the selection of condition indicators (spectral bands) has been performed without taking into account the specific dynamic characteristics of the Pelton turbine studied, but the typical bands used for this type of machinery.

The first condition indicator is the spectrum level of the turbine. The whole vibration spectrum (magnitude of the spectrum after applying the Fast Fourier Transform) is considered in order to build this set of indicators. According to the configuration of our monitoring system, the frequency range is 3-800 Hz with a resolution of 0.5 Hz, so there are 1595 lines in each spectrum. For each measuring point, the input pattern contains 1595 columns.

The second set encompasses the harmonics levels of the spectrum. The spectral bands of the rotating speed and its harmonics are excited by unbalance or misalignment of the rotor. Different types of damage can be reflected by changes in the harmonics of f_f [15,30,40]. The energy of a spectral band with a width of 2Hz and containing each harmonic is calculated according to Equation 2. Since the range of the spectrum is 3~800 Hz, 78 indicators are extracted from the raw data. Therefore, for one measurement point, the input contains 78 columns.

Traditional data-driven methods usually operate on the indicators without physical meaning. Generally, the routine process consists in extracting indicators and then reducing the dimension of the indicators' set by PCA [18,41]. Therefore, for the "spectrum level" and the "harmonic level" sets, which contain a lot of indicators with no clear physical meaning, the PCA technique is used to reduce the dimension.

Figure 13 shows the PCA result of the spectrum and harmonics indicators on measuring point A31-04. From the scree plot it can be seen that the first 18 components are able to explain more than 95% of the whole data contained in the "spectrum level" set of A31-04. Therefore, only these components are retained for this indicators' set. Similarly, 13 PCs are retained for "harmonic level" set of A31-04. The numbers of retained indicators of all measurement points using these two indicators' set are listed in Table 5.

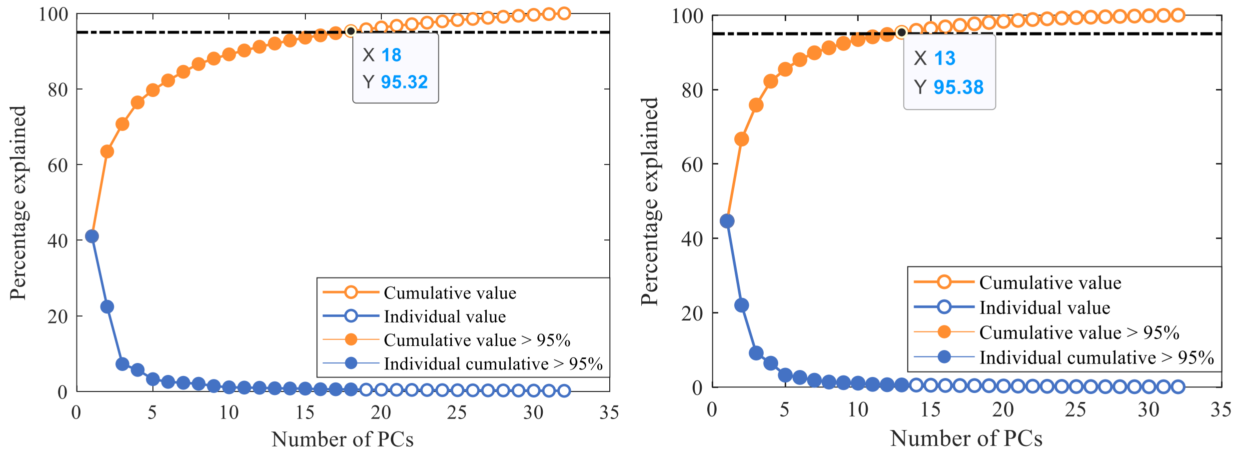


Figure 13 PCA scree plot of the "spectrum level" and "harmonic level" indicators' sets on point A31-04

Table 5 Numbers of retained indicators for each measurement point

Point	A11-04	A11-10	A14-04	A14-10	A31-04	A31-10	A34-04	A34-10	A35-04	A35-10
Spectrum level	15	17	15	12	18	21	16	17	17	18
Harmonics level	14	13	13	12	13	14	12	12	14	11

3.4. Definition of Artificial Damage Index (ADI)

The typical growth of a crack due to fatigue cycles is represented in Figure 14. According to [38], the growth of a defect is slow during the first stage and develops very rapidly after "High cycle fatigue onset". The development of the crack is divided into three stages: good condition, incipient damage and developed damage. This kind of evolution makes the detection of the incipient damage even more challenging as the crack size rapidly increases just before failure occurs. To evaluate the prediction performance of the different combinations of damage indicators, an Artificial Damage Index (*ADI*) is introduced. For the evolution of the *ADI* a power function, which is more progressive than the crack size evolution, is selected so that the sensitivity of the ANN for damage detection is increased. In equation 5, t is the normalized time with a range of 0 to 1, where t=0 is the moment when the machine starts the operation without any damage and t=1 is the time when failure occurs.

$$ADI = t^{1.85}, (t \in [0,1]) \tag{5}$$

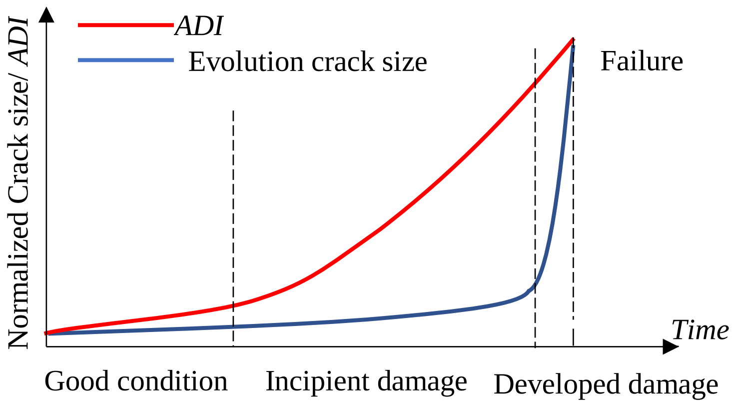


Figure 14 The curve of real crack size and the Artificial Damage Index

4. Results

The final set of indicators after applying the dimension reduction techniques (FA and PCA) are used to train the back-propagation neural networks (BPNN), whose output is the *ADI*. For each measuring point, the best perceptron number of the ANNs is optimized. Given there are only 33 samples (<50) in the database, leave-one-out cross validation (LOOCV) is applied to evaluate the prediction performance of different data sets. In LOOCV, each sample in turn is removed and the model is refitted using the remaining observations. The process is repeated again and again until every observation is predicted and predicted only once. The Root-Mean-Square Error (RMSE) of the prediction residuals are compared.

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}\left(ADI_p - ADI_0\right)^2} \tag{6}$$

Where ADI_p is the predicted damage index in each round of LOOCV and n equals to the total number of observations, which is 33 in our case. For each point, 50 trials have been carried out on the 3 groups of indicators respectively. Their *RMSE* have been calculated and are shown in Figure 15.

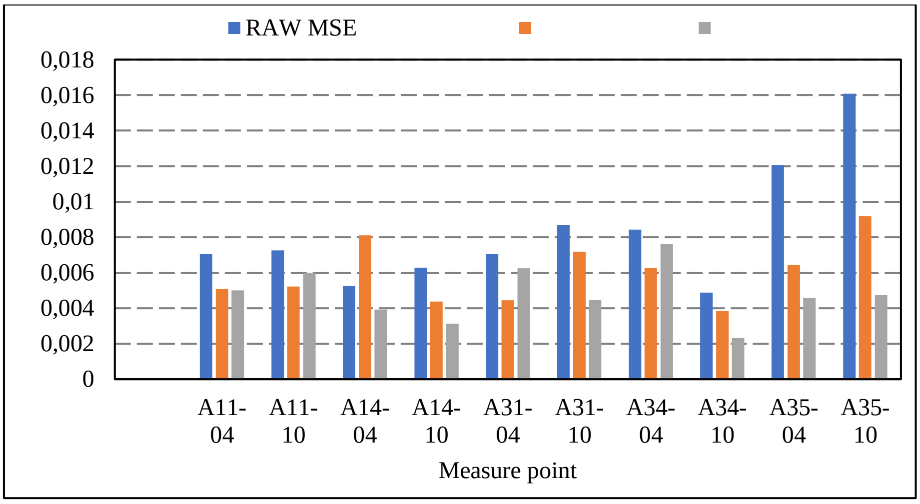


Figure 15 Prediction performance of the three indicators' set for each measurement point and averaged performance

It can be observed from the chart that the prediction result of the "spectrum level" set has the largest error in every point and the proposed set of indicators has the lowest error in most of the measurement points. The mean *RMSE* in the rightmost item in the chart indicates that the proposed condition indicators have the lowest error among the three data sets. Compared to the "spectrum level" and "harmonic level" indicators, the prediction error of new condition indicators has been decreased by 23.62% and 10.88% respectively. The most accurate point for the damage detection is the vertical position when the machine is working at 100% load (A34-10) and when the data set containing the new condition indicators is used. It can be concluded that the new condition indicators have a better correlation with the *ADI* for almost all the sensors analyzed (minimum RMSE). The reason lies in the characteristics of the especially selected spectral bands in the proposed set of indicators, which are more correlated to both the dynamic response of the structure and excitation forces. Thus, these indicators are more sensitive to detect the onset of an incipient damage.

Sensor A31 is located on the same bearing as A34, measuring vibrations in the horizontal direction. However, the prediction error is much higher than A34. This can be explained by the dynamic response of the structure: the weight of the runner makes the oil film thinner in the journal bearing, which makes the coherence between the bearing and the shaft higher on the vertical direction than in the horizontal direction. Another reason lies in the direction of the horizontal water jet, which causes too much noise on the measurement.

To assess the influence of dimension reduction by FA, the whole indicators in A34-10 before FA and after FA are used for predicting the *ADI*. Results are shown in Figure 16. It can be seen that the dimension reduction did not affect the *ADI* prediction and therefore the less important indicators according to FA technique could be removed.

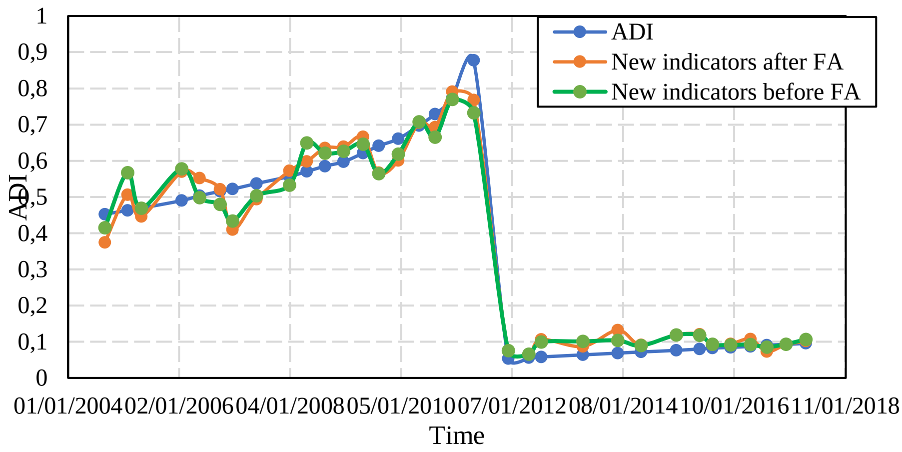


Figure 16 Predicted ADI by indicators before and after FA in point A34-10

5. Conclusions

In this paper, a new method to detect incipient damage in Pelton turbines has been presented. This has been attained by combining the expertise on the dynamic behavior of the machine with data-driven techniques including FA and ANN. The first step consists in extracting the new condition indicators from the vibration monitoring data according to the main excitations and the natural frequencies of the machine. After that, FA is carried out on the indicators: the ability of each indicator to take the damage information is quantified and the indicators which have low loading values to all of the common factors are eliminated from the data matrices. Finally, the selected indictors are used for training neural networks to form a damage detection model.

To verify the performance of the proposed method, an actual case of a Pelton turbine was studied. The studied machine suffered a failure where one bucket of the runner broke off during operation due to a deviated jet. The vibration data was recorded during normal operation, before and after failure took place and after the runner was repaired. An artificial damage index was defined according to the monitoring data and used as the output for the networks' training. As comparative groups, two traditional indicators' sets were introduced: a set containing the whole spectrum and a set containing the harmonics of the rotational speed. The dimensions of the comparative damage indicators were reduced by PCA.

The mean squared errors of each indicators' set show that the proposed new condition indicators' set have the best damage detection performance. For this new condition indicators' set the prediction error is decreased by 23.62% and 10.88% compared to the spectrum and harmonics indicators respectively. The measurement point A34 for the machine working at 100% rated output is selected as the best monitoring position. The indicators related to the rotor modes, runner modes, bucket passing frequency and rotating frequency have been determined by FA as the most important indicators for predicting the damage on the runner, which justify the proposed new condition indicators adopted in this paper. In conclusion, the final set of proposed indicators contains more relevant information and has a smaller dimension compared to the previous ones.

Acknowledgments

The authors want to acknowledge the XFLEX HYDRO project (EU H2020 No. 857832). Weiqiang Zhao would like to acknowledge the China Scholarship Council (CSC) for its grants. Alexandre Presas and David Valentin wants to acknowledge the Serra Hunter Program of Generalitat de Catalunya. The authors are grateful to the reviewers for their insightful comments and suggestions.

References

[1] A. Whiteman, J. Esparrago, S. Rueda, S. Elsayed, I. Arkhipova, RENEWABLE CAPACITY STATISTICS 2019, Abu Dhabi, 2019. www.irena.org/Publications.

[2] M. Egusquiza, E. Egusquiza, D. Valentin, C. Valero, A. Presas, Failure investigation of a Pelton turbine runner, Eng. Fail. Anal. 81 (2017) 234–244. doi:10.1016/j.engfailanal.2017.06.048.

[3] D. Ferreño, J.A. Álvarez, E. Ruiz, D. Méndez, L. Rodríguez, D. Hernández, Failure analysis of a Pelton turbine manufactured in soft martensitic stainless steel casting, Eng. Fail. Anal. 18 (2011) 256–270.

[4] J.C. Chávez, J.A. Valencia, G.A. Jaramillo, J.J. Coronado, S.A. Rodríguez, Failure analysis of a Pelton impeller, Eng. Fail. Anal. 48 (2015) 297–307.

[5] C. Vessaz, E. Jahanbakhsh, F. Avellan, Flow simulation of a Pelton bucket using finite volume particle method, in: IOP Conf. Ser. Earth Environ. Sci., IOP Publishing, 2014: p. 12003. doi:10.1088/1755-1315/22/1/012003.

[6] E. Egusquiza, C. Valero, D. Valentin, A. Presas, C.G. Rodriguez, Condition monitoring of pump-turbines. New challenges, Measurement. 67 (2015) 151–163. doi:10.1016/j.measurement.2015.01.004.

[7] ISO, ISO 20816-5:2018 Mechanical vibration -- Measurement and evaluation of machine vibration -- Part 5: Machine sets in hydraulic power generating and pump-storage plants, (2018) 60.

[8] D. Valentín, A. Presas, M. Egusquiza, C. Valero, E. Egusquiza, Transmission of High Frequency Vibrations in Rotating Systems. Application to Cavitation Detection in Hydraulic Turbines, Appl. Sci. 8 (2018) 451.

[9] D. Valentín, A. Presas, M. Bossio, M. Egusquiza, E. Egusquiza, C. Valero, Feasibility of Detecting Natural Frequencies of Hydraulic Turbines While in Operation, Using Strain Gauges, Sensors. 18 (2018) 174.

[10] D. Valentín, A. Presas, C. Valero, M. Egusquiza, E. Egusquiza, Detection of Hydraulic Phenomena in Francis Turbines with Different Sensors, Sensors. 19 (2019) 4053.

[11] W. Zhao, M. Egusquiza, C. Valero, D. Valentín, A. Presas, E. Egusquiza, On the Use of Artificial Neural Networks for Condition Monitoring of Pump-Turbines with Extended Operation, Measurement. (2020) 107952.

[12] M. Egusquiza, E. Egusquiza, C. Valero, A. Presas, D. Valentín, M. Bossio, Advanced condition monitoring of Pelton turbines, Measurement. 119 (2018) 46–55. doi:https://doi.org/10.1016/j.measurement.2018.01.030.

[13] A. Presas, Y. Luo, Z. Wang, B. Guo, Fatigue life estimation of Francis turbines based on experimental strain measurements: Review of the actual data and future trends, Renew. Sustain. Energy Rev. 102 (2019) 96–110. doi:10.1016/j.rser.2018.12.001.

[14] E. Egusquiza, C. Valero, A. Presas, X. Huang, A. Guardo, U. Seidel, Analysis of the dynamic response of pump-turbine impellers. Influence of the rotor, Mech. Syst. Signal Process. 68–69 (2016) 330–341. doi:https://doi.org/10.1016/j.ymssp.2015.05.034.

[15] G.P. Succi, H. Chin, Helicopter Hydraulic Pump Condition Monitoring Using Neural Net Analysis of the Vibration Signature, SAE Technical Paper, 1996.

[16] M. Schlechtingen, I.F. Santos, Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection, Mech. Syst. Signal Process. 25 (2011) 1849–1875.

[17] C. Zang, M. Imregun, Structural damage detection using artificial neural networks and measured FRF data reduced via principal component projection, J. Sound Vib. 242 (2001) 813–827.

[18] R.A. Saeed, A.N. Galybin, V. Popov, 3D fluid–structure modelling and vibration analysis for fault diagnosis of Francis turbine using multiple ANN and multiple ANFIS, Mech. Syst. Signal Process. 34 (2013) 259–276. doi:https://doi.org/10.1016/j.ymssp.2012.08.004.

[19] D.D. Kosambi, Statistics in function space, in: DD Kosambi, Springer, 2016: pp. 115–123.

[20] K. Pearson, On lines and planes of closest fit to systems of points in space (1901), Cited On. (1975) 26.

[21] H. Hotelling, Analysis of a complex of statistical variables into principal components., J. Educ. Psychol. 24 (1933) 417.

[22] I. Jolliffe, Principal component analysis, in: Int. Encycl. Stat. Sci., Springer, 2011: pp. 1094–1096.

[23] L.R. Fabrigar, D.T. Wegener, R.C. MacCallum, E.J. Strahan, Evaluating the use of exploratory factor analysis in psychological research., Psychol. Methods. 4 (1999) 272.

[24] C. Spearman, “ General Intelligence” Objectively Determined and Measured., (1961).

[25] J.C. Watson, Establishing evidence for internal structure using exploratory factor analysis, Meas. Eval. Couns. Dev. 50 (2017) 232–238.

[26] J. Baglin, Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR, Pract. Assessment, Res. Eval. 19 (2014) 2.

[27] M. Tripathi, S.K. Singal, Allocation of weights using factor analysis for development of a novel water quality index, Ecotoxicol. Environ. Saf. 183 (2019) 109510.

[28] R. Maskey, J. Fei, H.-O. Nguyen, Use of exploratory factor analysis in maritime research, Asian J. Shipp. Logist. 34 (2018) 91–111.

[29] E. Egusquiza, C. Valero, X. Huang, E. Jou, A. Guardo, C. Rodriguez, Failure investigation of a large pump-turbine runner, Eng. Fail. Anal. 23 (2012) 27–34. doi:10.1016/j.engfailanal.2012.01.012.

[30] E. Egusquiza, C. Valero, A. Estévez, A. Guardo, M. Coussirat, Failures due to ingested bodies in hydraulic turbines, Eng. Fail. Anal. 18 (2011) 464–473. doi:https://doi.org/10.1016/j.engfailanal.2010.09.039.

[31] M. Egusquiza Montagut, Study of the dynamic behavior of Pelton Turbines, Universitat Politècnica de Catalunya, 2020.

[32] M. Zhang, D. Valentin, C. Valero, M. Egusquiza, W. Zhao, Numerical Study on the Dynamic Behavior of a Francis Turbine Runner Model with a Crack, Energies. 11 (2018) 1630.

[33] H.H. Harman, Modern factor analysis, University of Chicago press, 1976.

[34] A.B. Costello, J. Osborne, Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis, Pract. Assessment, Res. Eval. 10 (2005) 7.

[35] C.C. Aggarwal, Neural Networks and Deep Learning: A Textbook, Springer International Publishing AG, Yorktown Heights, 2018.

[36] M. Gagnon, S.A. Tahan, P. Bocher, D. Thibault, The role of high cycle fatigue (HCF) onset in Francis runner reliability, in: IOP Conf. Ser. Earth Environ. Sci., IOP Publishing, 2012: p. 22005.

[37] X. Liu, A. Presas, Y. Luo, Z. Wang, Crack growth analysis and fatigue life estimation in the piston rod of a Kaplan hydro turbine, Fatigue Fract. Eng. Mater. Struct. 41 (2018) 2402–2417.

[38] A. Presas, Y. Luo, Z. Wang, B. Guo, Fatigue life estimation of Francis turbines based on experimental strain measurements: Review of the actual data and future trends, Renew. Sustain. Energy Rev. 102

(2019) 96–110.
[39] M. Egusquiza, C. Valero, D. Valentín, A. Presas, E. Egusquiza, Dynamic response of Pelton runners: Numerical and experimental analysis in prototypes, Renew. Energy. (2020).
[40] E.E. Estévez, E. UPC, Comportament dinàmic de màquines hidràuliques, Edicions UPC, Barcelona, 2003. http://hdl.handle.net/2099.3/36745.
[41] M. Ahmed, M. Baqqar, F. Gu, A.D. Ball, Fault detection and diagnosis using Principal Component Analysis of vibration data from a reciprocating compressor, in: Proc. 2012 UKACC Int. Conf. Control, IEEE, 2012: pp. 461–466.