FEATURE EXTRACTION FOR VIBRATION ANALYSIS OF CAVITATION IN KAPLAN WATER TURBINES

Esko Juuso
Control Engineering Laboratory, Department of Process and Environmental Engineering
P.O.Box 4300, FI-90014 University of Oulu, Finland
email: esko.juuso@oulu.fi

Sulo Lahdelma
Mechatronics and Machine Diagnostics Laboratory, Department of Mechanical Engineering
P.O.Box 4200, FI-90014 University of Oulu, Finland
email: sulo.lahdelma@oulu.fi

Pekka Vähäoja
Mechatronics and Machine Diagnostics Laboratory, Department of Mechanical Engineering
P.O.Box 4200, FI-90014 University of Oulu, Finland
email: pekka.vahaoja@oulu.fi

ABSTRACT
Several intelligent cavitation indicators obtained from vibration measurements have been compared in a Kaplan turbine. The indicators are based on the nonlinear scaling of features: one of the features is rms value and the other is either kurtosis or peak value. Indicators obtained from acceleration $x^{(2)}$ and higher derivatives $x^{(3)}$ and $x^{(4)}$ were tested by comparing the calculated indices with the sound of the recorded acceleration signals and analysing the signals with an oscilloscope in a wide power range. The results were compared in four frequency ranges with the knowledge-based cavitation index and previous studies. The indicators detect the normal operating conditions, which are free of cavitation, and also provide a clear indication of cavitation already at an early stage. The indices obtained from $x^{(4)}$ are the best alternative though also the index obtained from $x^{(3)}$ provides good results throughout the power range. Acceleration provided a good fit with the data but was less sensitive than higher derivatives. Automatic monitoring can be based on steps: detecting normal conditions, cavitation and the type of cavitation. The indicator also provides warnings of possible risk on short periods of cavitation. Uncertainties can be taken into account by extending the feature calculations and classification rules to fuzzy set systems.

KEYWORDS: Cavitation, vibration analysis, higher order derivatives, feature extraction and water turbine

1 INTRODUCTION
Condition monitoring provides a reliable and economical way of action for maintenance operations in modern industrial plants. The increasing number of measurement points and more demanding problems require automatic fault detection. Advanced signal processing methods expose failures earlier and provide information on suitable operating conditions for machines. Intelligent methods have been increasingly used in model-based fault diagnosis and intelligent analysers and provide various techniques for combining a large number of features.

Cavitation is extremely harmful to turbines, because it damages the surfaces of runners and flow channels. For instance, a revision of the runner may cause delays of several weeks in the operation of the turbine. The power ranges should be selected in a way that minimizes the
possibility of cavitation. For instance, if one wants to produce as high output powers as possible at flood periods, advantages and disadvantages can be estimated when the severity of cavitation is known at maximum power levels.

Figure 1. Kaplan water turbine.

Cavitation in water turbines has often been examined with standard $v_{\text{rms}}$ measurements in the frequency range 10 – 1000Hz or by selecting 1 Hz as the lower cut-off frequency. However, practical experiences have indicated that this analysis does not provide a sufficient picture of cavitation in the case of Kaplan turbines (Fig. 1). Higher order time derivatives have already been used to improve this analysis earlier [1].

There have been efforts to detect cavitation using vibration, pressure, acoustic emission or sound measurements. Cavitation and avoiding cavitation in water turbines has been investigated in some recent studies [2, 3, 4]. The measurement parameters traditionally used in vibration analysis are displacement, velocity and acceleration, i.e. $x^{(0)}$, $x^{(1)}$ and $x^{(2)}$. The parameters $x^{(3)}$ and $x^{(4)}$ are very suitable for condition monitoring when the ability to detect impacts is important [5, 6]. Lahdelma has used time derivatives of higher order than acceleration, such as $x^{(3)}$ and $x^{(4)}$, in various condition monitoring studies and these parameters are also very sensitive in detecting cavitation in water turbines as seen in references [1, 7].

Intelligent methods provide various techniques for combining a large number of features [8, 9]. Fault diagnosis and intelligent analysers are combined in model-based diagnostical process analysis (MDPA) [10]; the resulting systems can be used in various ways suitable for software sensors, risk analysis and the detection of sensor failures. Linguistic equations are designed for integrating knowledge and data in the development of non-linear multivariable systems for intelligent process analysis, process control, fault diagnosis and forecasting [11, 12]. An insight to the process operation is maintained since all the modules can be assessed on the basis of expert knowledge, and membership definitions relate measurements to appropriate linguistic terms on different operating areas [13].
Juuso and Lahdelma introduced in [14] a nonlinear scaling approach to features generated from vibration measurements and compared several cavitation indices constructed from these features in a water turbine. This paper continues the cavitation analysis by comparing different features and frequency ranges in the calculation of cavitation indices.

2 VIBRATION MEASUREMENTS

The turbine shown in Figure 1 has sleeve bearings and four blades and its rotation speed is 115 rpm. The turbine operates with a constant rotation speed and the output power is controlled by varying the volume flow rate of water through the turbine. Acceleration measurements were carried out vertically on the supporting bearings at 29 power levels varying from 1.5 to 59.4 MW. A Wilcoxon accelerometer model 726 with a permanent magnet was used in the measurements and signals were recorded using a Casio DA-7 16-bit DAT recorder in the frequency range 10 Hz to 20 kHz. The recorded acceleration signals were transferred to a computer using the LabVIEW 8.0 software and NI PCI-4472 24-bit data acquisition card. The sampling frequency in the data transfer was 12800 Hz, which is rather low, but sufficient in this case. The use of a higher sampling frequency would naturally increase the absolute values of different features but also increase sample file size. Certain features at varying frequency ranges were calculated from the stored acceleration signals using LabVIEW 8.0 software. The features were processed using the MATLAB software, version 7.3.

The numerical derivation and integration of the acceleration signals were performed with LabVIEW, and all the signals were filtered by means of a sixth order Butterworth bandpass filter. The frequency ranges were 10-1000 Hz, 10-2000 Hz, 10-3000 Hz and 10-4000 Hz. In this way, the time domain signals were generated at different power levels. An example signal \( x^{(4)} \) in the frequency range of 10-4000 Hz is presented in Figure 2. An analog differentiator/integrator was used in the previous study [14].

![Figure 2. Signal \( x^{(4)} \) at 2 MW power level in the frequency range 10-4000 Hz.](image)

The \( x^{(4)} \) signal at the 2 MW power level has strong, stochastic impacts (Figure 2). For instance, the signal at 10 MW has slightly more impacts, whereas the level of these impacts is lower. The signal levels at the cavitation-free power range are very low. Vibration levels grow when the power is increased above the clearly cavitation-free area, but the detected impacts are significantly weaker than in the 2 and 10 MW measurements. An example of signal \( x^{(4)} \) and corresponding acceleration signal \( x^{(2)} \) in the frequency range of 10-2000 Hz are shown in Figure 3. As seen, acceleration signals can reveal cavitation, but \( x^{(4)} \) is more sensitive.
In a previous study, strong cavitation was observed at power levels 2, 10 and 59.4 MW. There is a favourable, cavitation-free, operating range from 13 to 40 MW. For the low power range, i.e. cases of 1.5-1.8, 3-6, 8, 9, 11 and 12 MW, these short periods of cavitation were seen and heard from the acceleration signals as there are fairly strong impacts in the turbine. These impacts were much weaker than in the stronger cavitation cases and clearly stronger than in the case of 7 MW. For the high power range, i.e. cases of 45, 50, 56.5, 57 and 57.5 MW, the cavitation periods are seen and heard as a slightly higher noise level. The cases 58 and 58.1 MW have a continuously higher noise level. [14]

The cavitation indices based on the signals $x^{(3)}$ and $x^{(4)}$ presented in [14] efficiently indicated all these levels of cavitation. If all the index values $I_{c}^{(4)}$ above 40 MW are slightly reduced, also the cavitation index supports the fact that the favourable operating range extends to 50 MW. The index value for strong cavitation points at the low power range is emphasised by increasing the index values $I_{c}^{(4)}$ for the low power range. The resulting knowledge-based cavitation index $I_{c}^{*}$ is presented in Figure 4.

The indices $I_{c}^{(1)}$ based on the features of velocity do not indicate cavitation [14]. The rms and peak values obtained from velocity signals also have very low correlation with the knowledge-based cavitation index $I_{c}^{*}$ in all the frequency ranges. Therefore, velocity was dropped out from the detailed analysis. Other signals, $x^{(2)}$, $x^{(3)}$ and $x^{(4)}$, were analysed in each frequency range by rms values, kurtosis and peak values. Features for the frequency ranges 10-1000 Hz and 10-4000 Hz are presented in Figures 5 and 6, respectively. As the peak values are based on the highest three peaks in the discretized values, the three values may also originate from a single peak.
Figure 4. Knowledge-based cavitation index.

The kurtosis of each signal correlates with the index $I^*_C$ in the low power range: the strong cavitation at 2 MW is clearly detected. For the power range from 13 MW, kurtosis is close to value 3, which corresponds to a Gaussian signal, i.e. kurtosis does not give an indication of cavitation in the high power range. The spikes caused by cavitation are hidden in the signal since the noise level is increasing. However, the indication is achieved with the rms values, which includes the effects of spikes and noise. Both the features need to be combined in the power range 3 …12 MW: the cavitation point at 10 MW corresponds to high values in both the features. Only short periods of cavitation at 5 MW are needed to raise kurtosis when the signal levels are low, i.e. the rms value is low. An alternative feature for kurtosis is peak value, which has fairly similar changes in the low power range and small changes in the high power range.

Figure 5. Features of the signals $x^{(2)}$, $x^{(3)}$ and $x^{(4)}$ at different power levels in the frequency range 10-1000 Hz.

Widening the frequency range makes the features more sensitive in detecting cavitation at 10 and 59.4 MW. Also the absolute values of the features increase with the widening frequency range. The changes in kurtosis and peak values are different in different power ranges, e.g. the relative height of the spike becomes stronger at 10 MW and weaker at 5 MW. However, expanding the upper limit frequency from 2000 to 4000 Hz does not improve possibilities to observe cavitation at the power level of 2 MW. Thus the detection of cavitation does not require the use of very high upper limit frequencies. Even the upper limit frequency of 1000
Hz seems to be sufficient for detecting strong cavitation in this turbine. Model-based cavitation indices are needed for more detailed analysis.

![Figure 6](image1.png)

**Figure 6.** Features of the signals $x^{(2)}$, $x^{(3)}$ and $x^{(4)}$ at different power levels in the frequency range 10-4000 Hz.

Features were transformed to a linguistic scale from -2 to 2 by nonlinear scaling for the linguistic equation (LE) approach introduced by Juuso [11]. The scaling was done for the features in different frequency ranges by means of membership definitions, which for the features of the signal $x^{(4)}$ are shown in Figure 7. As the shape of the functions is quite similar, the features provide a good basis for the model-based analysis.

![Figure 7](image2.png)

**Figure 7.** Membership definitions of the features of a Kaplan water turbine for $x^{(4)}$: rms, kurtosis and peak values obtained for four frequency ranges (columns from the left: 10-1000 Hz, 10-2000 Hz, 10-3000 Hz, 10-4000 Hz).
4 MODELLING

Models for the cavitation index were developed in all four frequency ranges from the features obtained at ten power levels: 2, 3, 5, 8, 12, 25, 45, 57.5, 58.1 and 59.4 MW. The training set has to contain examples of cavitation, short-term cavitation and cavitation-free cases. Features of the frequency range 10-1000 Hz are presented in Figure 8. The same knowledge-based cavitation index IC * was used for all the frequency ranges (Figure 9 a). Six alternative variable groups shown in Figure 9 b were studied: power and features based on velocity were not excluded, variable 14 is the index IC *. The groups analysed for the signals x (2), x (3) and x (4) contain rms value and either the kurtosis or peak value of the signal.

![Figure 8. Features used in training in the frequency range 10-1000 Hz.](image)

(a) Knowledge-based cavitation index IC *.

![Figure 9. The goal of training and variable groups used in all the frequency ranges.](image)

(b) Variable groups for model alternatives.

The equations were generated for all the groups using linear regression for the scaled feature values. All the 24 models provide a good fit to both the train and test material. The example shown in Figure 10 has very high $R^2$ value: a considerable error only occurs at 7 MW power level (case 7), and another classification error at 57.5 MW power level (case 26). Also the model surfaces are quite similar: an increase of any input feature increases the cavitation index. There are two types of model surfaces (Figure 11), where the differences between frequency ranges are caused by the value ranges of the features as expected from the similarities of the membership definitions (Figure 7).
Figure 10. Testing results for the model based on the rms and peak values of the signal $x^{(4)}$ in the frequency range 10-3000 Hz.

Figure 11. Examples of model surfaces in the frequency ranges (a) 10-1000 Hz, and (b) 10-3000 Hz.

5 MONITORING

Cavitation, short term cavitation and cavitation-free operating conditions need to be detected in condition monitoring. Considerable differences between the models are clearly seen in the confusion matrices shown in Figure 12. The actual classes are denoted by $i = 1, 2$ and 3, and identified classes by $j = 1, 2$ and 3. Cavitation-free conditions (class 1) are reliably detected with features obtained in the frequency range 10-1000 Hz. The most difficult part is to find differences between cavitation (class 3) and short term cavitation (class 2). Although the acceleration features provided good fits to the train and test data, they are not sensitive enough for detecting the operating conditions (rows 1 and 2). The only successful result was obtained in the frequency range 10-4000 Hz with the model based on the rms and peak values (row 2, column 4). The features of the higher derivatives $x^{(3)}$ and $x^{(4)}$ have much better overall performance (rows 3...6). In the frequency ranges 10-1000 Hz and 10-2000 Hz, the results are quite similar for the signals $x^{(3)}$ and $x^{(4)}$. For wider frequency ranges, 10-3000 Hz and 10-4000 Hz, the features of the signal $x^{(4)}$ provide better results. The best performance was obtained by the rms and peak values of the signal $x^{(4)}$ in the frequency range 10-3000 Hz (row 6, column 3).

For the frequency ranges 10-1000 Hz and 10-2000 Hz, the combination of the rms value and the kurtosis is the best option for all the signals. For wider frequency ranges, 10-3000 Hz and 10-4000 Hz, the peak value is a better feature than the kurtosis. This is quite natural since kurtosis becomes less sensitive to impacts as noise increases. This model is closely related to the model used in [14] where peak values were calculated in a different way. Also the fraction of the values exceeding the normal range has clear similarities with the rms values.
Variation with time can be handled as uncertainty by presenting the indices as time-varying fuzzy numbers. The classification limits can also be considered fuzzy. The reasoning system will produce degrees of membership for different operating conditions. [14]

Figure 12. Confusion matrices $F(i,j)$, $i=1,2,3$ and $j=1,2,3$, for case detection with the features:
1. $[\text{rms}(x^{(2)}), \text{kurtosis}(x^{(2)})]$
2. $[\text{rms}(x^{(2)}), \text{p}(x^{(2)})]$
3. $[\text{rms}(x^{(3)}), \text{kurtosis}(x^{(3)})]$
4. $[\text{rms}(x^{(3)}), \text{p}(x^{(3)})]$
5. $[\text{rms}(x^{(4)}), \text{kurtosis}(x^{(4)})]$
6. $[\text{rms}(x^{(4)}), \text{p}(x^{(4)})]$

in four frequency ranges: 10-1000 Hz, 10-2000 Hz, 10-3000 Hz, 10-4000 Hz.

## 6 CONCLUSIONS

Features of velocity $x^{(1)}$, acceleration $x^{(2)}$ and higher derivatives $x^{(3)}$ and $x^{(4)}$ were compared in detecting cavitation. The features of velocity had a very low correlation with the knowledge-based cavitation index. Strong cavitation can be detected with the features of other signals in the low frequency range, even in the range 10-1000 Hz. Kurtosis combined with rms values was the best combination in the low frequency range, but widening the frequency range...
makes the peak values better than kurtosis. Several model-based indicators based on nonlinear scaling and linear equations provide a good fit to the cavitation index. However, only higher derivatives can be used in practical classification for cavitation, short term cavitation and cavitation-free operating conditions. The indices obtained from $x^{(4)}$ are the best alternative.

References

7. Lahdelma, S. Experiences of using higher derivatives in fault detection, Kunnossapito, 16 (2002) 9, pp. 29-32. (In Finnish)