MACHINE LEARNING FOR CONDITION MONITORING AND INNOVATION

N.H. Pontoppidan[1,a] *T. Lehn-Schiøler*[2] *K.B. Petersen*[3]

[1]Oticon, Rørtangvej 20, 3070 Snekkersten, Denmark
[2]Danske Bank, Holmens Kanal 2-12, 1092 København K, Denmark
[3]Laerdal Copenhagen, Njalsgade 19D, 2300 København, Denmark

ABSTRACT

This paper is a tribute to the late Professor Jan Larsen's work on machine learning for condition monitoring of large diesel engines. We present the ultrasound signals used to monitor the condition of the large engines and revisit two methods for estimating the condition of the engines from the ultrasound signals. Finally, we touch upon the increased importance of condition monitoring in large diesel engine industry today.

***Index Terms*—** Machine Learning, Condition Monitoring, Ultrasound

1. INTRODUCTION

The late professor Jan Larsen spent his long career on research and innovation with machine learning research in audio and related domains. In this paper we describe some of his work and advances based on ultrasound. From 1999, Jan Larsen worked on condition monitoring of large diesel engines with ultrasound signals in a number of research and innovation projects funded by Danish Academy of Technical Sciences, European Union, and industry partners [1, 2, 3, 4, 5]. Later from 2009, Jan Larsen's research interests related to diesel engines changed into predicting propulsion performance [6, 7, 8].

Large diesel engines are the most powerful combustion engines used for ship propulsion and power generation at secluded sites. The cylinder bore (diameter) ranges around 25 cm to almost 100 cm where the largest engines deliver more than 80 MW (corresponding to around 1000 cars with 100 HP) from one engine. The engines are used in the ships that transport goods around the globe, and here engine failure can cause huge monetary losses.

Traditional maintenance follow conservative time-based schemes, where specific parts get inspected and replaced at defined intervals that balances the active operation time against likelihood of failure. Condition monitoring enables a maintenance scheme based on the estimated condition, where early fault detection leads to inspection and replacement of parts beginning to approach worn out state. That means that costly unnecessary maintenance can be avoided and the time spend on maintenance can be minimized. Recent surveys of the maritime industry [9, 10] shows the innovation potential of Jan Larsen's work on condition monitoring.

[a] Corresponding authour: npon@eriksholm.com

2. DATA

When engines run, mechanical events like combustion, fuel injection, fluid and gas flows, impacts, frictions, and crack formation generate waves and depending on the frequency range and the sensor used to record it the waves can be look upon as vibration, sound, or ultrasound signals. The attenuation as function of distance is larger for ultrasound signals than for vibration signals and is therefore considered superior to vibration data for monitoring specific structures within an engine [11]. Figure 1 shows an example of an ultrasound signal from a large diesel engine, where the most dominant event, combustion, happens around 0^o when the mix of air and fuel ignites and the expansion initiates. For more details about the other events in the ultrasound signals see [12] and [13]. For further information about data and how it was obtained see [14], which describes the engine and the conditions of the engine during operation.

Data recorded from engines can be visualized in many different ways, some of the most useful representations for condition monitoring are visualized in Figure 2:
a) Plot of the waveform: This plot is simply a plot of one revolution of the raw measured ultrasound signal sampled in the time domain – (upper left).
b) Plot of the RMS signal: The RMS signal is the original waveform processed by the following steps: 1) A band pass filter, typically in the range 50-150 kHz, 2) Sum of squares in bins determined by the crank angle, 3) a logarithm – (upper right).
c) Color image of multiple RMS signals: A number of RMS plots are arranged side by side and given a color corresponding to its RMS value. The benefit of this is the possibility to get a visual inspection of many RMS signals at the same time – (bottom left).
d) Spectrogram: The spectrogram of one revolution reveals the frequency information over time by a fourier transformation in consecutive time windows. This provides a color im-

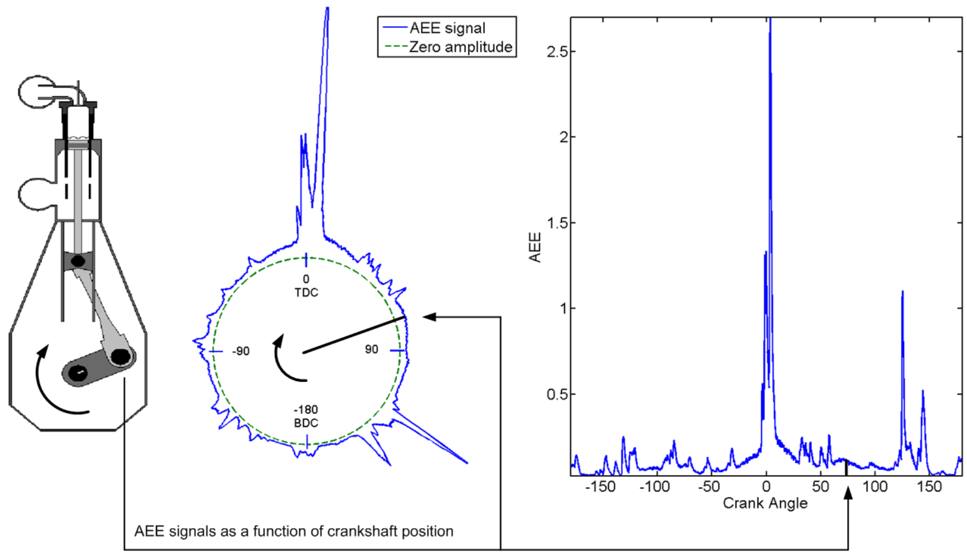
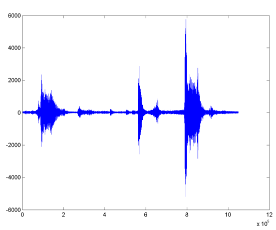
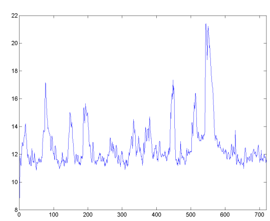


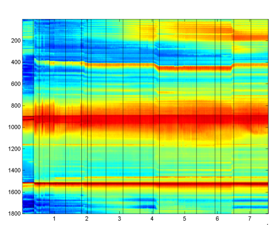
Fig. 1. An example of an ultrasound signal from one cycle on a diesel engine. From left to right: left: The engine with the crankshaft here at around 75^o, middle: the ultrasound signal plotted in a circular fashion, and right: the ultrasound signal plotted from -180^o to 180^o Adopted from [14].



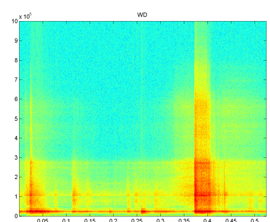
(a) Time domain



(b) RMS



(c) Multiple RMS



(d) Spectrogram

Fig. 2. Four visualizations of ultrasound signals.

age of the evolution in signal frequency – (bottom right).

3. METHODS

This section describes the methods for feature extraction and decision making.

The analysis uses an unsupervised approach, that is the change points must be detected without knowing examples of change. Change can occur on many time scales, sample to sample changes due to noise, changes on the time scale of hours or on the time scale of minutes, as is the case for changes in load or lubrication.

3.1. Mean Field ICA

In Independent Component Analysis (ICA) observed data vectors $\mathbf{x}_t$ are considered a result of underlying unobserved signals $\mathbf{s}_t$ which are mixed linearly by the matrix $\mathbf{A}$. Thus, it is assumed that $\mathbf{A}$ and $\mathbf{s}_t$ exists such that $\mathbf{x}_t = \mathbf{A}\mathbf{s}_t$ for all $t = 1, ..., T$ and when the unobserved signals $\mathbf{s}_t$ furthermore are assumed statistically independent and non-gaussian, the unobserved entities $\mathbf{A}$ and $\mathbf{s}_t$ can be estimated – see [15] or [16] for general introductions to ICA.

There are many different methods for estimating $\mathbf{A}$ and $\mathbf{s}_t$. In the so-called Mean Field ICA, the problem is approached using a two step approach where the matrix $\mathbf{A}$ is estimated using some estimated sources $\mathbf{s}_t$ followed by a step where the estimates of the sources $\mathbf{s}_t$ are updated using a fixed value of $\mathbf{A}$. The approach is motivated by methods from statistical physics dealing with the result of external magnetic fields and in a combination with Bayesian statistics, the signals $\mathbf{s}_t$ and the mixing matrix $\mathbf{A}$ can be restricted to be non-negative by choosing a non-negative prior – see [17] for more details on Mean Field ICA.

In relation to using ICA for condition monitoring of diesel engines, the observed signals $\mathbf{x}_t$ are assumed to be the mix of the underlying signals $\mathbf{s}_t$, each representing aspects of the engine which can easily be interpreted. In that sense, the estimated non-negative sources $\mathbf{s}_t$ are conjectured to be closely related to the actual condition of the engine.

3.2. Density model of the current state

The central idea of the density model of the current state is that one does not try to model all the states, but rather models the current state and identifies deviances from this state.

Density modeling is a very large scientific field and the methods described here are chosen for their simplicity and explanatory properties.

A certain part of data is used to build the density model, this data is defined as the *current* state, in this case a sliding window of 3000 samples was used. Then the probability that the next 100 points belong to the same distribution is calculated.

The 3000 rotations are represented by the PCA, ICA or NMF components $(\boldsymbol{z}_1 \ldots \boldsymbol{z}_T)$ most often 8 components are used. This data set represents the current state.

The density is build as a sum of multivariate Gaussian densities, each with the simplified covariance matrix $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, and centered on the data points $\mathbf{z}_t$ in the current state.

The probability density of a new point under this distribution is:

$$p(\boldsymbol{z}_n) = \sum_{t=1}^{T} \frac{1}{T} \mathcal{N}(\boldsymbol{z}_n | \mathbf{z}_t, \sigma^2\mathbf{I}).$$

A new data point, $\boldsymbol{z}_n$, can be evaluated and the density value, $p(\boldsymbol{z}_n)$, will have contributions from the densities of all data points in the current state.

Under this distribution data points $\boldsymbol{z}_n$ close to the current state have large $p(\boldsymbol{z}_n)$, while data points far from the normal state will have small values of $p(\boldsymbol{z}_n)$.

But, what is a large density value? To measure this, one can compute a so-called *Q-value*, which is related to the cumulative distribution function of the density values of the state. The Q-value is defined as

$$Q(\boldsymbol{z}_n) = Prob\{p < p(\boldsymbol{z}_n)\} \in [0\ 1].$$

That is, a Q-value $Q(\boldsymbol{z}_n) = 0.01$ means that only 1% of the data points in the current state had density values smaller than $p(\boldsymbol{z}_n)$.

Finally, the probability is calculated and is summarized for the 100 rotations to be evaluated against the new state.

4. RESULTS

First we look at what source separation offers in condition monitoring in terms of analyzing the contents of the data. A sample of 2000 ultrasound signals stacked as shown in Figure 2C was processed by Mean Field ICA [17] and the mixing matrix shown in Figure 3 and the source activations shown in Figure 4. Figure 3 shows the mixing matrix as four ultrasound signals that each describe the ultrasound signal during an engine revolution under a given condition. Referring to Chapter 2 in [14] the input data consisted of 2000 ultrasound signals recorded while the engine load increased from low to medium to high and while the lubricating oil system was disabled for a longer duration between revolution 180 and 1980. Signal 2 to Signal 4 in Figure 3 shows how the engine ultrasound signal looks under 3 different load conditions, while the last ultrasound signal is connected to the disabled lubricating oil system. This becomes even clearer when looking at Figure 4 showing the source activation for each of the four signals in the 2000 examples. Here Source 2 to Source 4 appear to be almost mutually exhaustive - corresponding well with the engine load being at 3 different levels whilst Source 1 increases from just before revolution 200 to right before revolution 2000 thus corresponding well with the period of disabled lubrication oil system.

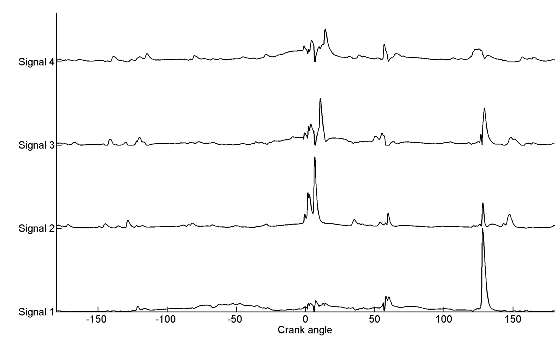


Fig. 3. Mixing matrix extracted with Mean Field ICA [17] using non-negative priors. Each of the four ultrasound signals shown from top to bottom represent the ultrasound signal during an engine revolution for four different operation conditions. Adopted from [14].

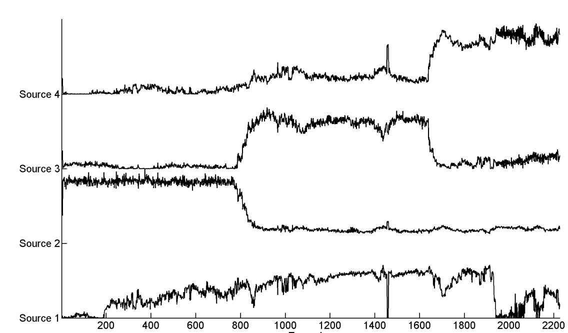


Fig. 4. Source activations extracted with Mean Field ICA [17] using non-negative priors. Each of the four source signals shown from top to bottom represent the strength of each ultrasound signal shown in Figure 3 as function of the revolution number. Adopted from [14].

Visual inspection of the elements of the mixing matrix reveals a striking feature of the Mean Field ICA outputs - all signals look like ordinary ultrasound signals. This is not the case for standard methods like Principal Component Analysis. However, even if the signals extracted with Mean Field ICA looked like real ultrasound signals, the advanced modeling of the ultrasound signals only led to slight improvements in the performance when identification of the impact of increased friction when compared to simpler models [5] .

In Figure 5 the result of the density estimation method is illustrated, using this simpler method change points are easily detected. The black curve shows the mean probability of 100 samples being similar to the previous 3000 samples a low score towards indicates that the ultrasound signal is changing and is thus indicative of either operational change or onset of a fault. With this method change points are easily detected, however, the changes could be caused by changes in engine load or other external events. Thus, to ensure that only relevant changes are marked as suspicious changes must be compared to the log from the engine control system. Alternatively changes across cylinders can be compared, if a change is present in all cylinders it is likely due to an external event, on the other hand, if it only occur in a single cylinder a warning should be raised.

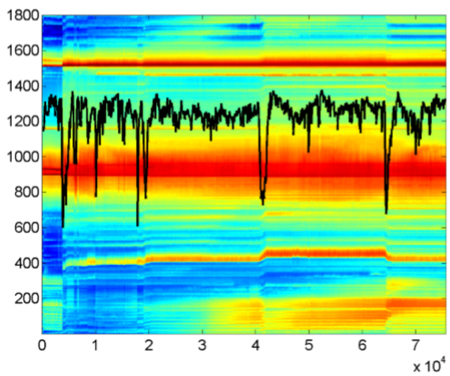


Fig. 5. Ultrasound RMS signals stacked horizontally under varying conditions (each revolution is represented by a vertical sequence of colors). The signals are colour coded such that blue is low intensity whilst cyan, yellow and red indicate increasingly higher intensity signals. The black curve shows the mean probability of 100 samples being similar to the previous 3000 samples.

The results shown in Figure 3-5 show how machine learning combined with other statistical modeling analyzes ultrasound signals from large diesel engines and separates out the partial ultrasound signals arising from the different loads as well as the fault. Moreover, the modeling of new revolutions based on the previous 3000 revolutions enables the detection of changes as they occur. Together, the two methods enable 1) detection of condition changes with the density modeling while 2) the Independent Component Analysis separates the ultrasound signal into the partials and activations. These partials and activations moreover reveal important information about the condition of the diesel engine.

5. IMPACT

Like many other industries, the large diesel engine industry is changing from offering products to offering services [10], e.g., from selling engines to selling propulsion as a service. In this new era, large diesel engine companies benefit from strong condition monitoring abilities, as it enables a competitive performance at a competitive cost, while the costs of preventive maintenance are minimized and operational availability maximized [18].

When Jan Larsen worked in this domain, condition monitoring was at best seen as a secondary feature, and definitely not one of the most competitive sales parameters. In hindsight, Prof. Larsen's research on utilizing machine learning in condition monitoring was visionary and ahead of its time and in fact anticipating the current needs of the large diesel engine industry.

Acknowledgements

We are forever grateful for the relationship we had with the late professor Jan Larsen. All research presented in this tribute paper was carried out together with Jan while first author was affiliated with Technical University of Denmark and the second and third author affiliated with the consulting company Epital.

6. REFERENCES

[1] T. Fog, L.K. Hansen, J. Larsen, H.S. Hansen, L.B. Madsen, P. Sørensen, E.R. Hansen, and P.S. Pedersen, "On condition monitoring of exhaust valves in marine diesel engines," in *Proceedings of the IEEE Workshop on Neural Networks for Signal Processing IX*, Y. H. Hu, J. Larsen, E. Wilson, and S. Douglas, Eds., Piscataway, New Jersey, 1999, IEEE, pp. 225–234.

[2] N. H. Pontoppidan and J. Larsen, "Unsupervised condition change detection in large diesel engines," in *2003 IEEE Workshop on Neural Networks for Signal Processing*, C. Molina, T. Adali, J. Larsen, M. Van Hulle, S. Douglas, and Jean Rouat, Eds., Piscataway, New Jersey, Sept. 2003, pp. 565–574, IEEE Press.

[3] N. H. Pontoppidan, J. Larsen, and T. Fog, "Independent component analysis for detection of condition changes in large diesels," in *COMADEM 2003*, Om P. Shrivastav, Bassim Al-Najjar, and Raj B.K.N. Rao, Eds. COMADEM International, 2003.

[4] N. H. Pontoppidan and J. Larsen, "Non-stationary condition monitoring through event alignment," in *IEEE*

Workshop on Machine Learning for Signal Processing, Piscataway, New Jersey, Sept. 2004, pp. 499–508, IEEE Press.

[5] Niels Henrik Pontoppidan, Sigurdur Sigurdsson, and Jan Larsen, “Condition monitoring with mean field independent components analysis,” *Mechanical Systems and Signal Processing*, vol. 19, no. 6, pp. 1337–1347, 2005.

[6] Benjamin Pjedsted Pedersen and Jan Larsen, “Prediction of full-scale propulsion power using artificial neural networks,” in *Proceedings of the 8th international conference on computer and IT applications in the maritime industries (COMPIT’09)*, 2009, pp. 10–12.

[7] Benjamin Pjedsted Pedersen and Jan Larsen, “Modeling of ship propulsion performance,” in *World Maritime Technology Conference WMTC2009*, 2009.

[8] Benjamin Pjedsted Pedersen and Jan Larsen, “Gaussian process regression for vessel performance monitoring,” in *12th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT 13)*, 2013.

[9] Krestine Mougaard, Line Maria Neugebauer, Adrià Garcia i Mateu, Jakob Axel Bejbro Andersen, Tim C. McAloone, Juliana Hsuan, and Thorkild Ahm, *Maritime Branch Analysis: A workbook in the PROTEUS series*, Technical University of Denmark (DTU), 2013.

[10] Krestine Mougaard, *A framework for conceptualisation of PSS solutions: On network-based development models*, Ph.D. thesis, Technical University of Denmark (DTU), 2016.

[11] G.D. Neill, S. Benzie, J.D. Gill, P.M. Sandford, E.R. Brown, J.A. Steel, and R.L. Reuben, *The relative merits of acoustic emission and acceleration monitoring for detection of bearing faults*, COMADEM, 1998.

[12] A.K. Frances, J.D. Gill, J.L. Reuben, and J.A. Steel, “A Study of the Variability of Acoustic Emission Signals from a Medium Size Marine Diesel Engine under Service Conditions,” in *Condition Monitoring and Diagnostic Engineering Management, Proceedings of the 16th International Congress*, Om P. Shrivastav, Bassim Al-Najjar, and Raj B.K.N. Rao, Eds. COMADEM International, 2003, pp. 503–512.

[13] R.M. Douglas, P. Nivesrangsan, A.I.F. Robertson, E.R. Brown, J.A. Steel, R.L. Reuben, and T.L. Fog, “Acoustic Emission as a Tool to Reveal Diesel Injector Performance,” in *Condition Monitoring and Diagnostic Engineering Management, Proceedings of the 17th International Congress*, Raj B.K.N. Rao, Barry E. Jones, and Roger I. Grosvenor, Eds., Birmingham, UK, 2004, COMADEM International, pp. 316–324, COMADEM International.

[14] Niels Henrik Pontoppidan, *Condition monitoring and management from acoustic emissions*, Ph.D. thesis, Technical University of Denmark, 2005.

[15] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent Component Analysis*, Wiley, 2001.

[16] Te-Won Lee, *Independent Component Analysis - Theory and Applications*, Kluwer Academic Publishers, 2010.

[17] Pedro Hojen-Sorensen, Ole Winther, and Lars Kai Hansen, “Mean field approaches to independent component analysis,” in *Neural Computation*, 2002, vol. 14, pp. 889–918.

[18] “Wärtsilä and Royal Caribbean extend their partnership to 2028,” *Navigator Magazine*, Oct. 2018.