Data driven case study of a wind turbine main-bearing failure

Elisha de Mello, Georgios Kampolis, Edward Hart, Daryl Hickey, Iain Dinwoodie, James Carroll, Rob Dwyer-Joyce and Ampea Boateng

1Department of Mechanical Engineering, The University of Sheffield, UK
2CDT in Wind and Marine Energy Systems and Structures, The University of Strathclyde, UK
3Wind Energy and Control Centre, EEE, The University of Strathclyde, UK
4Natural Power Consultants Ltd., The Greenhouse, Castle Douglas, UK
5ORE Catapult, Inovo Building, Glasgow, UK
E-mail: elxdemello1@sheffield.ac.uk

Abstract. This paper presents a data driven case study of two outer-race spalling faults in double-row spherical roller main-bearings, with the purpose of identifying key features and relevant measurements associated with this failure mode in wind turbine main-bearings. Supervisory data is analysed for one fault case and vibration data for the other. The aim of this work is to inform practitioners and assist in improving fault detection systems for this component.

1. Introduction

Wind energy is playing a key role in the decarbonisation of the power sector. In 2019 electricity production from wind power totaled 417 Terra Watt Hours, accounting for 15% of EU electricity consumption [1]. However, challenges still remain and concerted efforts are required to further reduce levelised cost of energy.

An important contribution to these cost reductions will come from amelioration of rates of catastrophic component failure through improved fault detection and predictive maintenance capabilities; cutting unplanned downtime and reducing costs associated with the replacement of damaged parts. Important advances in fault detection for wind turbine drivetrain components have been made in recent years [2, 3, 4, 5, 6], however, the main shaft bearing is a component which has seen less focus in the literature in this context [7] when compared to other drivetrain components. Furthermore, existing work which does consider fault detection and prognosis for main-bearings is often based on simulated and idealised data [8, 9, 10] or, in some cases, is based on scaled rotors whose inertias are orders of magnitude less than that of operational wind turbines [11]. Studies that do employ real world data (e.g. [12, 13]) can also be limited by the data types and number of failures they have access to, meaning proposed detection techniques are difficult to validate. Even in studies reporting promising main-bearing fault detection performance [13, 14], it is not clear whether employed data was ascertained to be the most appropriate for the task at hand, or simply what was available. Therefore, in order to support improvements in main-bearing fault detection capabilities, the current work looks to
study data from two examples of outer-raceway main-bearing failures (supervisory data from one, vibration data from the other) in order to demonstrate which aspects of this data can be associated with the known fault type. In addition, the study seeks to inform practitioners as to the most promising features for inclusion in detection algorithms. As such, focus will be on the data itself, and observable trends as a fault progresses, rather than the algorithms which might then make use of such data. Operating conditions for main-bearings are generally high-load and low-speed, making it also pertinent to ask whether classical fault indicators, commonly used in high-speed applications, are also present here. Note, while this manuscript was under review, two valuable additions to the main-bearing fault detection literature have been published [15, 16]. Employed data was from two sources: that collected from the supervisory control and data acquisition (SCADA) system, and vibration measurements of 100s duration at 8kHz using an accelerometer affixed to the main-bearing casing. The SCADA data consists of 10-min means and standard deviations of wind turbine operational and environmental variables, including: main-bearing temperature, rotor speed, wind speed and power. Note, due to hardware issues main-bearing temperature and vibration data was not available for the same turbine, the SCADA data analysed here was instead obtained from an identical turbine with the same fault type. The turbines' rated power is $\sim$2MW.

2. SCADA data analysis

SCADA data was analysed for 16 months before failure of the main-bearing outer raceway. Starting from first principles, it was assumed that the presence of damage within the bearing would lead to elevated temperatures at each operating point. However, wind turbine operation is highly variable, so even for healthy main-bearings time variations in temperature will be present. Main-bearing temperature ($T_{MB}$) was therefore plotted against key operational variables in order to help delineate between natural variability and damage induced effects. Figures 1-3 show these operational plots for rotor speed, power and wind speed respectively. Each includes data from month 4 (1 year before failure) and month 16 (1 month before failure). The expected temperature elevations are clearly present.

$$T_{adj} = T_{MB} - T_{amb}. \tag{1}$$
However, it was found that the use of $T_{\text{amb}}$ only increased the scatter present. A correlation analysis indicated why, since $T_{\text{MB}}$ and $T_{\text{amb}}$ were found to be completely uncorrelated. This suggests that, for this turbine at least, the nacelle generates its own independent micro-climate possibly due to thermal insulation or active temperature control.

In Figures 1-3, while some co-dependency is apparent between considered variables, there is also significant scatter which will, inevitably, reduce the sensitivity of fault detection algorithms which seek to monitor and detect changes in main-bearing temperature relationships. The lack of structure, seen in above plots, is perhaps somewhat puzzling given the clear physical link which exists between $T_{\text{MB}}$ and the other variables, since both load and speed would be expected to influence frictional effects fairly directly. Further analysis of the data was therefore undertaken in order to seek an explanation.

2.1. Thermal inertia and system response
To better understand the underlying causal relationships, time series of SCADA data values for $T_{\text{MB}}$ and the other variables were considered. Here, rotor speed values will be used to illustrate findings, with similar conclusions found for the other variables.

Figure 4 shows example concurrent $T_{\text{MB}}$ and rotor speed time series (scaled to be zero mean and unit variance - referred to here as “normalised intensity” time series). In this form, the reason for an observed lack of structure becomes clear. Changes in $T_{\text{MB}}$ can be seen to lag behind changes in rotor speed. This lag is caused by thermal inertia present in the system, in which the rate of accumulation and dissipation of heat is dictated by a range of factors such as material heat capacity, physical size, surface area and frictional surface geometry.

Relative to changes in rotor speed, the presence of thermal inertia elicits two main effects when comparing these time series. The first is a clear smoothing effect, where higher frequency variations present in rotor speed do not manifest in the temperature signal. The second is a lag between signals, whereby any sustained change in rotor speed only gradually generates an associated change in temperature values. Since the relationship between rotor speed and main-bearing temperature has this added level of complexity, the lack of structure observed in joint plots becomes unsurprising. This result then begs the following question: can rotor speed
values, \( \Omega \), be transformed into a new speed variable, \( \hat{\Omega} \), such that inertial effects are accounted for and the relationship between speed and temperature strengthened? An answer in the affirmative suggests an improved variable on which learning algorithms might be based or, alternatively, helps inform as to appropriate learning features for inclusion at the fault prediction stage.

2.2. Transforming rotor speed to account for inertia

Inertia-like effects can be synthetically incorporated into a transformed version of rotor speed time series via the following process:

(i) Application of a moving average across the time series. Moving averages of different forms, such as simple [17], auto-regressive integrated moving average (ARIMA) models [18, 19] and exponentially weighted moving average models (EWMA) [19] have been shown in the literature to help improve causal relationships in this type of context. Here a simple, unweighted, trailing moving average across \( N \) 10-min bins was applied; this is a simple example of a finite impulse response filter.

(ii) Application of a time translation, \( \tau \), to the output of step (i) to remove lag. \( \tau \) is chosen such that the correlation between \( \hat{\Omega} \) and \( T_{MB} \) is maximised.

The combined effect of these steps generates the transformed rotor speed time series, for example see Figures 5 and 6 for a given number of averaging bins \((N)\). It is clear that the relationship between \( \hat{\Omega} \) and temperature is now much stronger.

![Figure 5. Time series before preprocessing. The effects of thermal inertia are clearly present.](image1)

![Figure 6. Time series after preprocessing rotor speed using an averaging window of 22 and time-shift of 8.](image2)

This same idea can be extended so as to consider optimal values of both \( N \) and \( \tau \), again such that correlation between the resulting transformed rotor speed and \( T_{MB} \) is maximised. As such, the following cost function was defined,

\[
\rho_{\text{opt}}(N) = \max_{\tau} \rho \left( \hat{\Omega}(N, \tau), T_{MB} \right).
\]

(2)

\( \rho_{\text{opt}} \) returns the maximum correlation \((\rho)\) possible when a time shift \((\tau)\) is applied to the moving average processed rotor speed signal, where the averaging is over \( N \) bins. The associated time shift which achieves this maximum, \( \tau_{\text{opt}} \), is therefore the optimal lag for that choice of \( N \). Using \( \rho_{\text{opt}} \) to investigate optimal parameters across the available data, it was found that the outlined process results in significant increases in signal correlation for all values of \( N \), with correlations generally increasing from values as low as 0.69 to those as high as 0.95. With no averaging, \( N = 1 \), a large value of \( \tau \) is required, with \( \tau_{\text{opt}} \) corresponding to between 1.5 and 3
hours. However, the overall correlation between signals was found to be maximum when a large averaging window, \( N = 40 \text{–} 60 \) (~6.5-10 hours), was used in conjunction with no time shift, \( \tau_{\text{opt}} = 0 \). This last result is due to the fact that the application of a non-centred moving average also introduces some amount of lag into the resulting signal, with results indicating that at the point of maximum correlation the smoothing transform alone is enough to remove signal lag.

Figure 7 shows the \( \hat{\Omega}, T_{MB} \) relationship after applying the globally optimal transformation, as described above, for data from a year prior to failure (month 4) and one month prior to failure (month 16). When compared to Figure 1, the effects of pre-processing are clear, with much reduced scatter and a stronger relationship evident between the two variables. There remains a discernible increase in temperature values from month 4 to month 16, especially at higher rotor speeds where the overlap between datasets has been all but removed.

A key question at this stage is whether by transforming \( \Omega \rightarrow \hat{\Omega} \), the impending failure becomes easier to detect. This question was considered by comparing the shift seen in lines of best fit relative to the magnitude of residuals from those lines. More formally, for lines of best fit \( L_4 \) and \( L_{16} \) (written in vector form) on original data, and denoting those fitted to processed data as \( \hat{L}_4 \) and \( \hat{L}_{16} \); the relative shift in each case if defined as,

\[
\Delta T = \frac{\text{mean}(|L_{16} - L_4|)}{\sigma} \quad (3)
\]

\[
\hat{\Delta} T = \frac{\text{mean}(|\hat{L}_{16} - \hat{L}_4|)}{\hat{\sigma}} \quad (4)
\]

where \( \sigma \) denotes the average of original data residual standard deviations from lines \( L_4 \) and \( L_{16} \) respectively, and similarly for \( \hat{\sigma} \) with respect to \( \hat{L}_4 \) and \( \hat{L}_{16} \) on transformed data. With these definitions in place, it was found that the relative shift \( \hat{\Delta} T \) after transforming rotor speed values was around 40% larger than that obtained using the original data \( \Delta T \), indicating a significant improvement in sensitivity which could in turn allow for earlier or more reliable detection of impending main-bearing failures. In practise, more sophisticated methods than straight line fits would be used to detect faults, however, the analysis presented here highlights the potential benefits of either pre-processing as outlined, or ensuring such effects are accounted for in learning methodologies.

While correlation can be maximised with a large enough averaging window and no lag, it is also worth considering the possible information made available when a shorter averaging window
is applied, such that non-zero optimal lag values (for the chosen value of $N$) are obtained. Considering the degrading internal conditions of a failing main-bearing, which in turn lead to higher frictional forces, and so the higher temperatures observed in Figures 1-3, it would be reasonable to expect rougher surfaces and the presence of metallic debris to also effect the rate at which temperature responds to changes in speed (or load). The lag applied above can in fact be interpreted as an estimate of the rate at which bearing temperature responds to changes in rotor speed, an estimate learned from the data itself. Optimal lag values were therefore investigated, using intermediate values of $N$, in order to ascertain whether a reduction in $\tau_{\text{opt}}$ values tends to occur as the bearing moves closer to failure. On average this was found to be the case; for example, Figure 8 shows distributions of daily optimal lag values from the first and last five months of data using an averaging window of $N = 10$. Dashed lines show the average of optimal lag values in each case. Both with respect to the distributions themselves and their average values, a shift towards the left is evident for the data taken from closer to failure. This in turn implies that optimal lag might itself constitute a useful learning feature to be tracked and fed to fault detection algorithms.

The transformation process outlined in this section can also be applied to power and wind speed signals with similar results. Overall the highest correlations were to found exist in the rotor speed case.

3. Vibration data analysis
Vibration data was analysed for a turbine of the same type and power-rating, also with an outer-race fault. Note, due to hardware issues required SCADA and vibration data was not available for the same turbine, hence data from a similar turbine with the same fault type was investigated instead.

Vibration based fault detection is a well established field, and one in which changes in signal characteristics associated with given fault types are well understood. As with other approaches to fault detection, in the case of wind turbine main-bearings there are few studies which investigate vibration signal characteristics as the bearing progresses to failure, although the technique is commonly and successfully used by industry for main-bearing fault detection. The present section considers vibration signal characteristics from a number of healthy and one faulty main-bearing between 18 and 1 months to failure. As previously, the focus is on the data itself and, in this case, whether classical signs of failure are also present for this main-bearing fault example. Vibration signal characteristics will be considered in both frequency and time domains.

Figure 9 (a) shows a power spectral density (PSD) plot for the failed bearing (black) around 1 month before turbine shutdown. Also included are a range of healthy main-bearing PSDs from two different turbines (green and yellow) for comparison. Power spectra were estimated using Welch’s method with a Hanning window and no overlap. Cumulative energy spectra, obtained by integrating the square of PSD values, are also shown for the faulty case and one of the healthy cases. Energy values have been normalised, and so move monotonically from $-\infty$ to 0 on the log scale. Note also that frequency values are normalised by the mean rotational frequency, $f_r$, of the low speed shaft, measured during each vibration sample.

Rotor harmonics are clearly visible in all signals at 1P (likely due to blade imbalance) and 3P, as would be expected. Unique to the faulty case are three distinct peaks between 14 and 45 $f/f_r$. Such peaks can often be used to identify a fault location by comparing frequencies of identified peaks with those corresponding to repeated interactions between bearing internal surfaces. Standard formulae exist such that these characteristic frequencies can be determined from bearing geometry and rotational speed alone [20, 21, 22, 23]. In the current case, the additional peaks appearing in the faulty signal correspond to the first three harmonics (1st, 2nd and 3rd) of the ball-pass-frequency on the outer raceway (BPFO) for this particular main-bearing, thus corresponding to the known fault type of the failure in question. Also apparent
between healthy and faulty signals is a markedly raised noise floor.

The normalised cumulative energy spectra associated with these signals provide useful guidance as to which features should be considered significant. In the healthy case, almost all signal energy is shown to be concentrated in the fundamental and 3rd harmonic (3P) of rotational speed, indicated by the normalised energy converging to 0 (on the log scale) at, or just after, the 3P peak. In the faulty case, however, the total signal energy is spread across a much larger range of frequencies, only converging to 0 at around 100 $f/f_r$. Along with energy ‘steps’ at 1P and 3P, the three BPFO harmonics can be seen to also contain enough energy to visibly increase cumulative energy values. Curiously, the peak at around 95 $f/f_r$, which is visible but not significant (energy wise) in the healthy cases, becomes significant in the faulty case. The nature of this peak is still under investigation, the current hypothesis is that this is in fact a gearbox generated signal transmitted to the main-bearing via the low speed shaft. The energy content of this peak close to main-bearing failure indicates a possible interaction whereby additional loads/harmonics may be being transmitted to the gearbox as the main-bearing fails. However, this is currently only conjecture, more work is needed to conclusively account for this.

Figure 9. Vibration data results showing (a) power and (normalised) cumulative energy spectra for healthy and faulty bearings (b) cumulative energy spectra as damage progresses (c) time-domain signal statistics over time. Note, in (b) the green and black curves are the same as those shown in (a). The red curve is from faulty main-bearing data at an intermediate damage stage.
signal component. The shift in normalised cumulative energy observed in Figure 9 (a) was found to hold in general as the fault progressed. As illustrated by Figure 9 (b), which shows an intermediate damage state energy spectra from the faulty main-bearing, fault progression was found to be accompanied by the normalised energy curve moving down and to the right. It can also be seen in Figure 9 (b) that the intermediate fault data (red) has a dominant first BPFO harmonic, whereas closer to failure (black) the energy contributions of BPFO 1-3 are more uniform.

Cumulative energy characteristics observed in the current study indicate a degradation path where the shape of the consecutive curves results in a reduction of the overall area under the curve. This is similar to the evaluation of performance of a supervised classifier with the area under the curve (AUC) metric of a Receiving Operating Characteristic (ROC) curve (see for example [24, 25, 26]). The summarising property of AUC is potentially apt to depict, in a single value, the state of the dispersion of energy in the spectrum. As such, it can be used along with other summary statistics to compare successive values for sustained changes (trends) which may indicate a progression towards failure. This implies that the AUC of cumulative energy spectra for main-bearing vibrations signals could potentially prove a useful feature to include in fault detection algorithms. In this case, the shift of the normalised cumulative energy curve observed here would manifest as a reducing value of area over time.

Finally, time domain statistics of the measured vibration signals were analysed with a view to identifying any obvious trends. Figure 9 (c) shows chronological values of vibration signal variance, room-mean-square (RMS), peak-to-peak ratio (P2P), peak-to-RMS ratio (P2R) and signal-to-noise (SNR) ratio. For the number of measurements available, no discernible trends in the above quantities are apparent as the bearing moves towards failure, except perhaps in SNR where there is a possible downward trend over time. The relatively constant values seen for variance, RMS, P2P and P2R over time indicate it is likely not the case that natural variations (i.e. noise) in these quantities are masking underlying trends. At this stage, not enough data is available to say whether the observed SNR reductions are statistically significant and more work will be undertaken to better determine if this is the case. The P2P results seen here are in contrast to those of [12], where this statistic was found to be a useful indicator of a main-bearing fault. The fault type occurring in this previous work is not given, hence it could be the case that different time-domain features act as indicators for different fault types. It should also be noted that, in [12], a significantly larger dataset of P2P values was analysed. It could therefore be the case that with more data a similar trend would manifest here; however, if this were the case one would expect to see more variability in the P2P values shown in Figure 9.

4. Discussion
In this study, data from an outer race main-bearing failure was analysed. Results of the SCADA data analysis (Section 2) indicated that, as expected and in agreement with other studies [19], temperatures were observed to increase as the bearing moved closer to failure, reflecting increased frictional effects in the system. However, some subtleties associated with the data were also investigated, with findings indicating that thermal inertia effects should be considered when developing fault detection models. It was further demonstrated that improved relationships between key variables can be obtained by pre-processing measured data to artificially incorporate inertial effects. Such pre-processing procedures might therefore be incorporated into fault detection algorithms in order to improve sensitivities. Alternatively, results could be interpreted as providing insight as to the types of features necessary for inclusion in learning models (removing the need for the data itself to be manipulated outwith the learning procedure); explicitly, features should be included that allow for inertial type effects to be accounted for, such as smoothed measurements and lags of measured values. The choice between pre-processing or feature-level inclusion of these effects will largely depend on sophistication of
the chosen fault detection methodology.

It is also worth noting that while average based smoothing and time shifting of signals has been commonly used to ‘clean’ data prior to learning in the past, the current study indicates that this step could provide more than just improved input-output correlations. As shown in Section 2.2 optimal lags were found to reduce in mean value closer to failure, indicating they may form a useful additional learning feature in their own right. More generally, this could be extended to the use of other measures of system inertia for monitoring of main-bearing health. This suggest that, whilst main-bearing temperature itself is certainly a key indicator of system health, other measurable properties are present which shouldn’t be overlooked.

Similarly, in the vibration data case, signal characteristics were found to display classical features as the bearing moves to failure. In particular, harmonics associated with the known fault type were clearly visible. Additionally, the cumulative signal energy spectra and its associated area-under-the-curve metric was identified as another single variable feature on which fault detection might be based.

5. Conclusions
In this study, supervisory and vibration datasets were analysed with regards to their fault detection potential on an example of a wind turbine outer raceway fault. Classical signs of failure were present in both cases, with additional effects also investigated. It was found that thermal inertia present in the system warranted specific consideration, with changes in inertia associated time-lags potentially offering a new learning variable. Vibration data analysis showed that cumulative energy spectra might also provide a single variable on which sustained changes might be used to indicate a developing fault.

Acknowledgements
This work forms part of Projects AMBERS led by E. Hart. E. de Mello is funded through the Powertrain Research Hub (PTRH), co-funded by the Offshore Renewable Energy Catapult. G. Kampolis is funded by the EPSRC (EP/LO16680/1) and Natural Power. E. Hart is funded by a Brunel Fellowship from the Royal Commission for the Exhibition of 1851.

References
[1] “Wind Energy in Europe in 2019”. Wind Europe, 2019, retrieved 2021-01-25. https://windeurope.org/about-wind/statistics/european/wind-energy-in-europe-in-2019/
[2] M. D. Machado de Azevedo, A. M. Araújo and N. Bouchonneau. Renewable and Sustainable Energy Reviews, 56, 368-379, 2016, https://doi.org/10.1016/j.rser.2015.11.032
[3] A. Steteo, F. Dinmohammadi, X. Zhao et al. Renew. Energy, 133:620-635, 2019. https://doi.org/10.1016/j.renene.2018.10.047
[4] R. Moeini, M. Entezami, M. Ratkovac et al. Wind Engineering, 43(5):539-555, 2019. https://doi.org/10.1177/0309524X18807028
[5] A. Zaher, S. D. J. McArthur, D. G. Infield and Y. Patel. Wind Energy 12(6):574–593, 2009. https://doi.org/10.1002/we.319
[6] J. Carroll, S. Koukoura, A. McDonald, et al. Wind Energy, 22:360–375, 2019. https://doi.org/10.1002/we.2290
[7] E. Hart, B. Clarke, G. Nicholas, et al. Wind Energ. Sci., 5:105–124, 2020. https://doi.org/10.5194/wes-5-105-2020
[8] Y. Qu, C. Chen and B. Zhou. Advanced Materials Research. 308-310:1264-1268, 2011. https://doi.org/10.4028/www.scientific.net/AMR.308-310.1264
[9] M. Ghane, A. R. Nejad, M. Blanke, Z. Gao and T. Moan. J. Phys. Conf. Ser., 753, 052017, 2016. https://doi.org/10.1088/1742-6596/753/5/052017
[10] M. Hamadache and D. Lee. 16th International Conference on Control, Automation and Systems (ICCAS), Gyeongju, Korea (South), 1579-1584, 2016. https://doi.org/10.1109/ICCAS.2016.7832512

Advancing Main-BEaRing Science for wind and tidal turbines.
[11] J. Wang, Y. Peng, W. Qiao, and J. Hudgins. IEEE Transactions on Industry Applications. 53(3):3029-3038, 2017. https://doi.org/10.1109/TIA.2017.2650142

[12] R. Zimroz, W. Bartelmus, T. Barszcz and J. Urbanek. Key Engineering Materials. 518:437-444, 2012. https://doi.org/10.4028/www.scientific.net/KEM.518.437

[13] Z. Zhang. Open Journal of Applied Sciences, 8, 211-225, 2018. https://doi.org/10.4236/ojapps.2018.86018

[14] Z. Y. Zhang and K. S. Wang. Adv. Manuf. 2, 70-78, 2014. https://doi.org/10.1007/s40436-014-0061-6

[15] M. Beretta, A. Julian, J. Sepulveda, J. Cusidó, and O. Porro. Sensors, 21(4),1512, 2021. https://doi.org/10.3390/s21041512

[16] A. Encalada-Dávila, B. Puruncajas, C. Tutivén and Y. Vidal. Sensors, 21(6), 2228, 2021. https://doi.org/10.3390/s21062228

[17] A. Kusiak and A. Verma. Renewable Energy, 48:110-116, 2012. https://doi.org/10.1016/j.renene.2012.04.020

[18] P. Chen, T. Pedersen, B. Bak-Jensen, Z. Chen. IEEE Transactions on Power Systems. 25(2):667-676, 2010. https://doi.org/10.1109/TPWRS.2009.2033277

[19] Y. Zhao, D. Li, A. Dong, D. Kang, Q. Lv, L. Shang. Energies, 10(8), 1210, 2017. https://doi.org/10.3390/en10081210

[20] R. B. Randall and J. Antoni. Mechanical Systems and Signal Processing, 25(2):485-520, 2011. https://doi.org/10.1016/j.ymssp.2010.07.017

[21] D. S. Chandra and Y. S. Rao. J Fail. Anal. and Preven., 19:1144–1152, 2019. https://doi.org/10.1007/s11668-019-00712-z

[22] J. I. Taylor. The Vibration Analysis Handbook, Vibration Consultants, ISBN-13: 9780964051720, 2003.

[23] W. Teng, R. Jiang, X. Ding, Y. Liu and Z. Ma. Shock and Vibration. 2378435, 2016. https://doi.org/10.1155/2016/10155/2016/2378435

[24] A. P. Bradley, Pattern Recognition, 30(7):1145-1159, 1997. https://doi.org/10.1016/s0031-3203(96)00142-2

[25] D. J. Hand and R. J. Till. Machine Learning 45:171-186, 2001. https://doi.org/10.1023/A:1010920819831

[26] T. Fawcett, Pattern Recognition Letters, 27(8):861-874, 2006. https://doi.org/10.1016/j.patcog.2005.10.010