Twelfth International Congress
on Sound and Vibration

A CASE STUDY OF AUTOREGRESSIVE MODELLING AND ORDER SELECTION FOR A DRY VACUUM PUMP

Suguna Thanagasundram, Yanhui Feng and Fernando Soares Schlindwein

University of Leicester, Department of Engineering
University Road,
LE1 7HA Leicester, UK
st89@le.ac.uk

Abstract

Oil-free dry-vacuum pumps are prevalent in semiconductor plants worldwide today. The popularity of such pumps has grown over the years but little research has been done in the area of condition-monitoring of dry-vacuum pumps. Fault detection of one such pump through the use of AutoRegressive (AR) modelling technique and spectral analysis of vibration and acoustic data has been studied. The testing environment is a 5-stage Roots-and-Claw dry-vacuum pump. Usage of spectral analysis for fault prediction in real-time has been conservative due to concerns over large processing requirements especially when large sample sizes and high sampling frequencies are used. In this study it is shown how such concerns can be allayed, to a large extent, by AR modelling, as the AR method has enhanced resolution capabilities compared to the FFT technique even when small sample sizes are used and requires a sampling rate just slightly above Nyquist rate to give good parameter estimates. The AR spectra of both the vibration and acoustic data produced are highly correlated and the frequencies of the maximum peaks are found at the pump's shaft rotational speed and multiples of it. The disadvantage of the AR method is that optimum model order is not known a priori. It was desired to keep the model order low as smaller orders translate to smaller processing requirements for spectral estimation. Several methods of order selection criteria such as AIC (Akaike Information Criterion), FPE (Final Prediction Error), MDL (Minimum Description Length), CAT (Criterion Autoregressive Transfer-function) and the more recent FSIC (Finite Sample Information Criterion) were investigated to find the true order. All the criteria selected approximately the same order. For the vibration data, initial results show that the minimum order required can be as low as 25 and for 90% of the frames the model order does not exceed 45.

INTRODUCTION

Dry vacuum pumps are positive displacement mechanical rotary pumps that attain a vacuum without the use of lubricants in the pumping chamber. The particular dry vacuum pump that was used in the study is a BOC Edwards IGX modular multistage pump that has one stage of roots and four stages of claws. The study of this kind of pump has become important as the dry vacuum pump is especially the pump of choice in semiconductor and cleanroom environments [1] where corrosive gases are abundant, temperatures are high and effluents resulting from the processes can cause seizure of the pumps. The more traditional oil sealed rotary pumps have certain disadvantages in such conditions and were found to break down often in such hostile and aggressive environments even with expensive modifications such as use of filters, traps, forced lubrication and use of inert corrosive resistant materials for the construction of the pump housing. With the advent of oil-free pumping technology, dry vacuum pumps have become an effective and reliable alternative for such applications [2]. Dry vacuum pumps require no internal lubrication and have non-contacting rotors to pump process the vapors. They have no rubbing components to wear and since they do not use oil, there is no oil back migration and reduced atmospheric pollution [3]. The dry vacuum pumps can also easily handle the toxic vapors since they run hot and corrosion only takes place in presence of moisture.

The design of dry vacuum pumps has remained relatively the same over the last ten years, more emphasis is being employed in self-diagnosis capabilities of the health of the pump [4]. Pump failure can contribute to the significant loss of valuable products eg. loss of wafer batches with value in excess of $100,000+ range. Customers can be affected by the inconvenience of down time. Hence early detection of incipient faults has a primary customer requirement. Today in most wafer fabrication plants, vacuum pumps are fitted with integrated sensors that acquire various signals from the pump and these signals are processed by microprocessors to monitor the condition of the pumps. Two key parameters that can be studied, as indicators of the health of pump are the vibration and noise signals. Early symptoms of pump degradation and malfunction can be predicted from these signals. The acquired vibration and acoustic data in time domain are often complex and important features necessary for fault detection are often not obvious in the raw signatures. Spectral analysis of the data is necessary to convert the information contained in the data from a time domain representation to a frequency domain representation. The power spectrum estimation is normally done using the classical approach of FFT (Fast Fourier Transform) but the main limitation of this method in the field of machine condition monitoring is its poor frequency resolution[5] when applied to short lengths of signal. The frequency resolution of FFT based frequency is inversely related to the number of data samples used. The higher the number of samples used, the better is the frequency resolution achieved. However, in practical applications there is always a upper limit to the number of samples to use for spectral analysis as a bigger sample sizes translate to bigger memory buffers to hold the data as well as longer processing times for each frame, inherently affecting the performance if a real time spectral analysis tool is to be implemented. Also in some situations, only short lengths of signal are available. For our application, it was desired to have a frequency resolution Δf of 6.1 Hz as it known prior that the fault frequencies were tightly specified. This shortcoming of the FFT method can be overcome the parametric AR modelling technique which has increased spectral resolution capabilities and hence applicability to short data lengths. Unlike the FFT technique for spectral estimation, resolution is not improved in the AR model by oversampling the signal. A sample frequency of slightly more than twice the Nyquist rate is all that is required to give good parameter estimates and to capture the basic behaviour of the system.

Parametric modeling has been employed by [6] in fault diagnosis studies in low speed machinery. The effects of noise on the vibration signal and optimum vibration signal length have also been investigated using the AR model [7],[8]. Determination of the model order is a fundamental task in the parametric estimation The model order is not known a priori and some experimentation with different orders is required before the right order can be selected for the given finite record of signal of length *N* samples (frame). If the model order is too low, unmodeled dynamics will result in insufficient detail in the spectra leading to nonconfident fault diagnosis. If model order is too high, then the spectral estimate is too peaky and spectral line splitting might occur [9].A study has been carried out considering some issues of the AR model order selection[10].From this study, it has been concluded that overestimating the model order is better than underestimating it. Also it has been identified that using a smaller number of samples is more likely to underestimate the model order.

HARDWARE AND DATA ACQUISITION

To capture the vibration and sound signals, speed of the IGX pump was set at 100 Hz with a 0 mbar loading factor. Accelerometers were mounted radially on the dry vacuum pump near the high vacuum end. A microphone was also attached in the vicinity to acquire the sound signals. Both signals were suitably pre-filtered with in-house built 8^{th} order elliptic anti-aliasing filters with cutoff frequencies of 10 *kHz* and 20 *kHz* respectively. The analogue to digital conversion was performed with a 16- bit NI 6034E ADC card. The sampling rate was set to 50 *kHz* and samples of *N*=8192 were acquired. The data was then downsampled to 2 *kHz* as it was known that the fault frequencies lie in the range from 0-1kHz. The size of the downsampled frames are *N*=328.

AR MODELLING AND ESTIMATION ALGORITHMS

An AR time series of order *p* has the current value of the series $x[n]$ expressed as a linear function of *p* previous values plus an error term. $e[n]$. $e[n]$ is a purely stationary white noise signal with zero mean and finite variance σ^2.

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + e[n] \tag{1}$$

The AR series is defined as a stationary process because the poles of the AR transfer function, which are the roots of the characteristic polynomial a_k all lie within the init circle. It is a necessary condition that the sum of the a_k parameters must be less than unity in magnitude as $\left(\left| a_1 + a_2 + a_3 + \ldots\ldots\ldots\ldots a_p \right| < 1 \right)$ to avoid non-stationarity. The power spectrum can be derived from the a_k coefficients of the AR model, and is defined in (2) where $P_{AR}(f)$ is the AR power spectral estimation, T is the sample period, a_k are the AR coefficients, p is the number of a_k parameters or order of the AR model and σ^2 is the variance of the signal.

$$P_{AR}(f) = \frac{\sigma^2 T}{\left| 1 + \sum_{k=1}^{p} a_k e^{-j2\pi fkT} \right|^2} \qquad (2)$$

The four methods for the estimation of the AR parameters are [5]: Yule-Walker method, Burg method, covariance method and modified covariance method. It has been shown by Broersen [11] that the true order of the AR signal depends not only on the characteristics of the AR process and but also on the method of parameter estimation used. The solution used in this study is the Yule-Walker method that uses Levinson-Durbin recursion on the autocorrelation matrix to find the AR coefficients[5]. The Yule-Walker estimation method produces a biased estimate of the residual variance however it has been shown the effect of bias on order selection is negligible[12]. The Yule-Walker estimation method was chosen for computation of the a_k parameters as it has been proven by [12] that asymptotic criteria can produce a maximum followed by a second-order minimum which can be lower than the minimum at the true model order that should be selected when used with the Burg, covariance and modified covariance estimation methods.

ORDER SELECTION CRITERIA

Many methods of order selection criteria exist for order estimation. The true AR order of a given signal is unknown and has to determined via the application of these criteria. The most established of these include the Final Prediction Error (FPE)[13] and Akaike Information Criterion(AIC)[14] developed by Akaike which are classified as asymptotic information criteria. There have been reported to be many variants of the AIC method but the one used for this investigation is the original one with a penalty factor of 2 .The other two criteria tested was Rissanen's Minimum Description Length (MDL) [15] estimator and Parzen's Criterion AR Transfer (CAT*)* Function[16]. All these criteria are designed to reduce the probability of underfit at the cost of overfit [11]. A cost function term in each of them penalises for the use of extra AR parameters above the true order. For all the order estimation criteria stated,

N is defined to be the sample size, k the model order and σ^2 the variance of the prediction error for the given model order. The equations for the order selection criteria are given below:

$$FPE(k) = \frac{N+k+1}{N-k-1}\sigma(k)^2 \tag{3}$$

$$AIC(k) = \ln(\sigma(k)^2) + (2k+1)/N \tag{4}$$

$$MDL(k) = \sigma(k)^2\left(1+\left(\frac{p+1}{N}\right)\ln(N)\right) \tag{5}$$

$$CAT(k) = \left[\frac{1}{N}\sum_{j=1}^{k}\frac{N-j}{N\sigma(j)^2}\right] - \frac{N-k}{N\sigma(k)^2} \tag{6}$$

Basically all these criteria are functions of the finite variance of the prediction error σ^2. The variance of prediction error is defined as such:

$$\sigma(k)^2 = \frac{1}{N-p}\sum_{n=p+1}^{N}\left\{x(n) + \sum_{i=1}^{p} a_i x(n-i)\right\}^2 \tag{7}$$

The optimal model order is the one that minimizes the above criterion equations as a function of order. The performance of a more recent method of order selection criteria, Finite Information Criterion (FIC)[12] was also investigated in this study. This method has been stated to be more accurate than the asymptotic criteria for a finite number of samples as it has been shown by Broersen [12] that the asymptotic criteria do not take into account the distinction that exists in practice between the different estimation methods. Samples are called finite if $N < \infty$ and p/N is greater than a small value such as 0.1.The FIC has been proposed to perform better than the asymptotic criteria when the ratio of the number of samples used to the model order used in the investigation is large. Also this order selection criteria calculates the variance coefficients depending on the method of estimation used hence this method is meant to be more robust and applicable then the other order selection methods. The FIC criterion is defined (8) where v_i is the finite sample variance coefficient for the Yule-Walker method.

$$v_i = \frac{N-i}{\{N(N+2\}}$$

$$FIC(k) = \ln\left(\sigma(k)^2\right) + \alpha\sum_{i=1}^{p} v_i \tag{8}$$

RESULTS AND DISCUSSION

An assessment of the abilities in estimating the true optimal order of the AR process when applied to frames of vibration and sound signals of fixed sample sizes was investigated. The study consists of two main parts. In the first part, the “true orders” for describing 428 frames of N=328 samples of vibration signal using the five criteria were evaluated. A maximum order, K, was chosen. In our case this was fixed to be $K=60$, as it was expected that the optimum order would not exceed this value for the chosen data set. Variance $\sigma(k)^2$ of the prediction error was computed scanning orders from $0 \le k \le K$ using the Yule-Walker estimation method. Then for each specific criterion, the criterion values are calculated for the range $0 \le k \le K$ using (3), (4), (5), (6) and (8). The optimum order is independently selected for each criterion as the minimum criterion value for each method. The steps are repeated for all the frames in the signal as a Monte Carlo simulation. The distribution and range of model orders for each of these criteria as order occurrence versus model order was then plotted as histograms. The order with the highest occurrence is identified as the “true order”, M for each order selection criteria. The behaviour of the five criteria in identifying the true order is analysed. Ideally, all the five criteria should have suggested the same true order. However this is not the case as the results show.

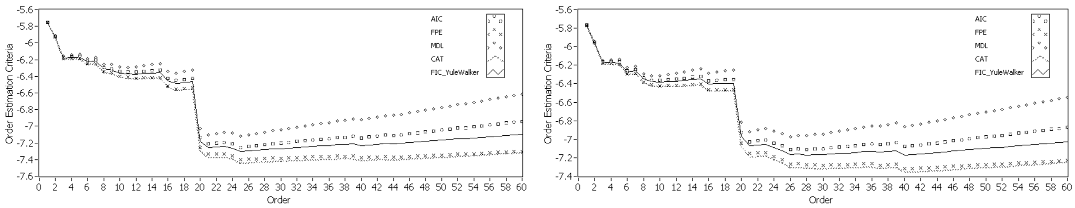

Figure 1 – (Left) Order selection criteria given as a function of model order using the Yule-Walker estimation method using a particular vibration data set with a sample size of N=328. Here K = 60 and the optimum order M=25 for all the order selection criteria (Right) Same simulations as used in (a) but using another data frame (sample size also N=328). For this particular frame, the optimum order M selected for AIC(k)=MDL(k)=FIC(k)=25.But FPE(k) and CAT(k) chose a slightly higher order of 41.

Figure 1(left) shows the plot of how the variance of the prediction error for every model order p has been transformed into order criteria *AIC(k), FPE(k), MDL(k),CAT(k)* and *FIC(k)* for orders $0 \le k \le K$. It can be seen for low model orders, all the criteria have close values. A significant finding is that all the criteria have a local minimum occurring at an order around 25.Hence for this particular frame, the best order M is equivalent to 25 for the given AR signal and sample size. For higher model orders, the criterion values deviate. For some of the frames, FPE and CAT criteria chose higher true orders. This phenomenon can be explained if one looks closely at Figure 1(right). There was a kink occurring at order around 41.The penalty factor in built in the other criteria had a monotonically increasing functions. So there weren’t affected very much by this kink and selected order 25 as the local minimum. But this effect wasn’t true for FPE and CAT in which case the penalty seems to be constant or slowly increasing. Since the optimum order was earlier defined to be minimum criterion value in the range $0 \le k \le K$, these two criteria indeed select the next best local minimum as the true order though this was not supposed to be the case. At this point the influence the choice of the maximum order K has on the order selection should be questioned. If it had been selected somewhere around 40, this problem could have been avoided and all the criteria would have given similar results. But this was not the case in this example as the maximum order *K* was fixed to be 60 for the simulations. We can now see how sensitive asymptotic criteria like FPE and CAT can be to the maximum order selected. The performance of these two criteria has also degraded because a finite number of samples were used. In fact it is known that FIC was the only criteria designed to work well independent of the choice of the maximum order used.

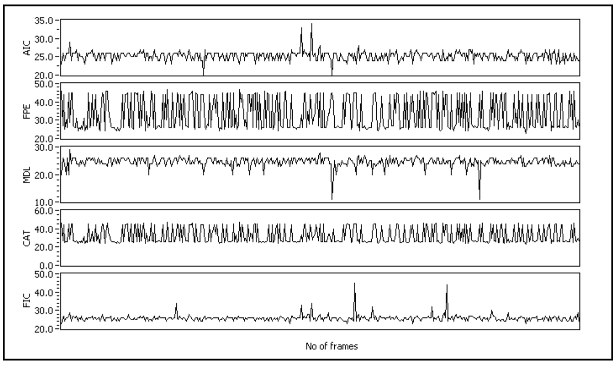

Figure 2 – Analysis of the statistics of the order of AR model for the various order estimation criteria using428 frames of vibration signal each with a frame size of N=328.

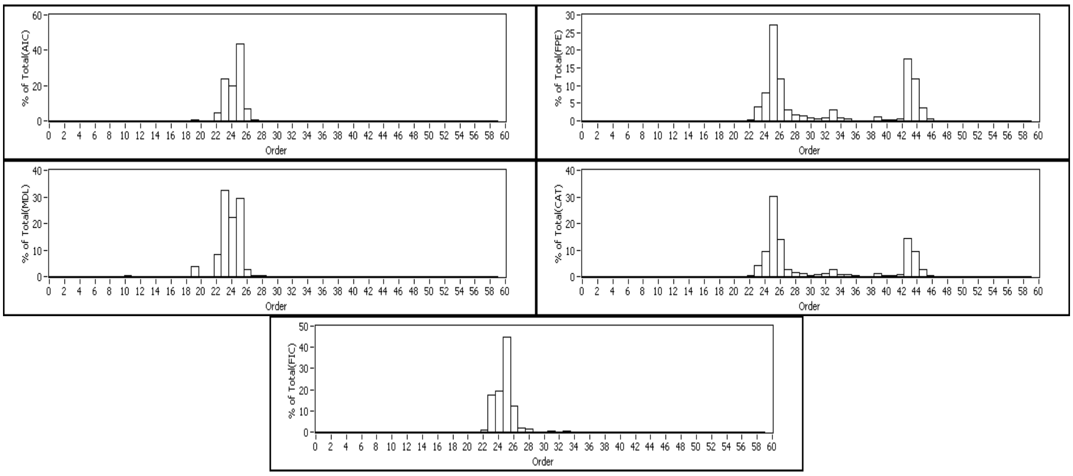

Figure 3 – 428 Frames Vibration Signal of Length N=328.

From Figure 2, it can be confirmed that FIC criteria is the best criterion for order selection in this case and it was found to outperform the other criteria when the sample size was small. The FIC criterion has performed considerably well and selected the true order of 25 for most of the frames. AIC and MDL have almost identical performance except at some points where the selected order differs. The performance of FPE and CAT criteria was poor. There is always a split between the orders selected. It is either 25 or 45.This can be explained by the earlier discussion for Figure 1.The occurrence of orders as histograms is given in Figure 3.The same trend of results can be observed here statistically. The distribution for each of the above criteria in selecting the true order is, in that order, is 42%, 27%, 23%, 30% and 46%. The highest order selected does not exceed 45 for in fact more than 90% of the samples. For the MDL criterion, the highest order selected is 23 but there is a clustering of orders around 23,24 and 25.The splitting behaviour of the orders can also be observed for the FPE and CAT criteria.

Verification of true order

From the previous investigation, true order *M* should for this particular vibration signal can be ascertained to be 25. Then a typical frame of data with the order 25 is selected and 25 a_k coefficients are generated. Sets of 1000 simulated series representing the AR process of various lengths *N*= 300, 500 and 800 samples are generated .The first 200 samples in each realization are discarded to minimise transient effects. Then the order selection criteria are again applied to the 1000 simulated AR processes to test their performance. The effect of increasing the sample size on the performance is also investigated.

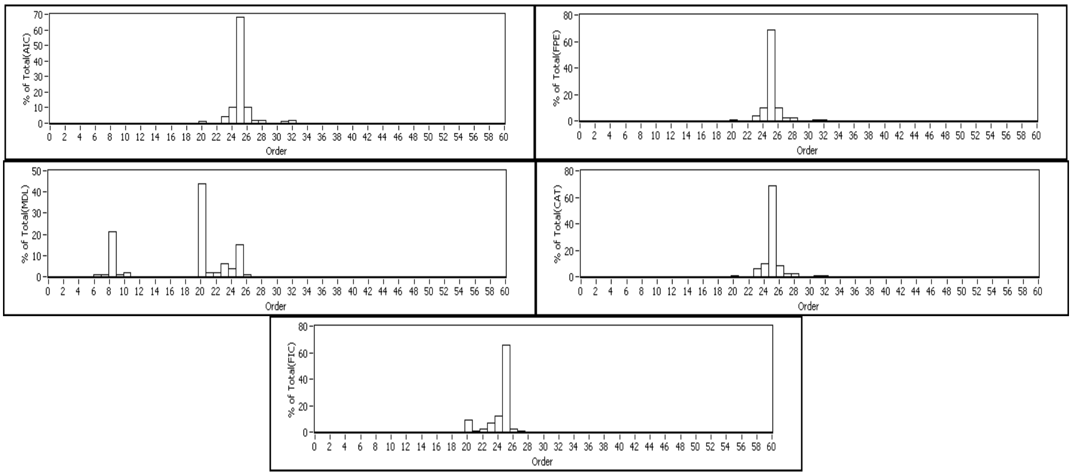

Figure 4 – Verification of order prediction criteria using a true AR process

In all cases the histograms have a peak corresponding to the true original order *M=25* except in the case of MDL. MDL has estimated a lower model order of 20.This is expected of the MDL criterion as it belongs to a class of consistent criteria which tend to select under-fitted models. From the spread of orders, it is seen that overestimation of order is negligible in all cases. Underestimation of order does occur but tends to minimize as sample size increases. This is alright as over-fitting is less problematic than under-fitting. Performance is improved in the MDL criteria as expected as the sample size increases(Figure 5).

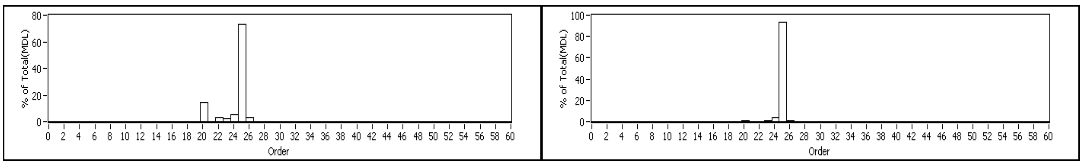

Figure 5 – Effect of increasing sample size for MDL criterion (Left N=500, Right N=800)

Comparison with the FFT method

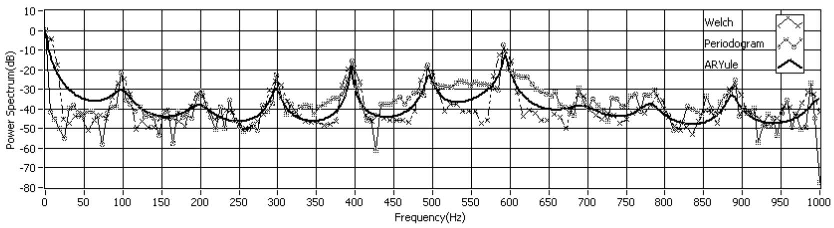

Figure 6 – Comparison of the power spectra produced by the FFT methods such as Welch and Periodogram and the parametric method of AR modelling on vibration data of sample size N=328 samples and true order determined M=25.

Figure 6 shows the frequency spectra of the vibration signal obtained using both the Fourier and AR techniques when the pump was rotating at 100Hz. The AR technique has produced far smoother and better spectrum than the Fourier technique. The main peaks required for fault diagnosis are clearly visible in AR spectra. The shaft rotational frequency of 100 Hz and its harmonics can be clearly seen. The weakness of the FFT techniques clearly becomes apparent in this example .The FFT spectra have poor resolution because a small sample size is used. The tests are repeated for the sound signal.

CONCLUSION

One of the main advantages of the use of the AR modelling for spectral estimation for fault condition monitoring is it superior resolution capabilities. The use of the AR model allows higher resolution achievable over the FFT technique especially with a smaller sampling rate and short data records. Its only disadvantage is that its order is not known and has to be determined via order selection criteria. In this study, it has been shown clearly how the true order can be accurately determined. The findings of this study can provide useful insight on the use of order selection criteria and on the use of the AR model for spectral estimation for fault condition monitoring in future researches. All the criteria perform well in determining the true order. However there is no significant gain in differentiating between these criteria unless when only a small number of samples are available as the performance of the criteria is dependent on sample size used. The order selection criteria are originally designed in such a way that larger orders are less preferable. The spirit of parsimony is against the selection of higher orders and the simpler the AR model is, the better it is. Order selection criteria can give inconsistent results if applied to short frames of data due to the availability of a small number of samples and in such cases care has to be taken to avoid inaccurate results. Asymptotic criteria like AIC, FPE, MDL and CAT perform better only when the sample size is large. If the number of observations used is small, the ability of the order selection criteria in identifying the correct autoregressive order may be affected. In such cases, higher orders may be selected due to the occurrence of a second local minimum. Which order selection criterion to use also depends on intended application.

REFERENCES

1. Troup, A.P. and D. Turrell, *Dry pumps operating under harsh conditions in the semiconductor industry.* Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, 1989. **7**(3): p. 2381-2386.
2. Troup, A.P. and N.T.M. Dennis, *Six years of ``dry pumping'': A review of experience and issues.* Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, 1991. **9**(3): p. 2048-2052.
3. Wong, W., et al., *An evaluation of the composition of the residual atmosphere above a commercial dry pump.* Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, 1988. **6**(3): p. 1183-1186.
4. Duval, P., *Will tomorrow's high-vacuum pumps be universal or highly specialized?* Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, 1987. **5**(4): p. 2546-2551.
5. Kay, S.M., S.L. Marple, and Jr., *Spectrum analysis-a modern perspective.* Proceedings of the IEEE, 1981. **69**(11): p. 1380-1419.
6. Mechefske, C.K. and J. Mathew, *Parametric spectral estimation to detect and diagnose faults in low speed rolling element bearings: Preliminary investigations.* Mechanical Systems and Signal Processing, 1993. **7**(1): p. 1-12.
7. Mechefske, C.K., *Machine condition monitoring: part 1 - optimum vibration signal lengths.* British Journal of Non-Destructive Testing, 1993a. **35**(9): p. 503-507.
8. Mechefske, C.K., *Machine condition monitoring: part 2 - the effects of noise in the vibration signal.* British Journal of Non-Destructive Testing, 1993b. **35**(10): p. 574-579.
9. Fougere, P.F., E.J. Zawalick, and H.R. Radoski, *Spontaneous line splitting in maximum entropy power spectrum analysis.* Physics of The Earth and Planetary Interiors, 1976. **12**(2-3): p. 201-207.
10. Schlindwein, F.S. and D.H. Evans, *Selection of the order of autoregressive models for spectral analysis of doppler ultrasound signals.* Ultrasound in Medicine & Biology, 1990. **16**(1): p. 81-91.
11. Broersen, P.M.T., *Finite Sample Criteria for Autoregressive Order Selection.* IEEE Transactions on Signal Processing, 2000. **48**(12): p. 3550-3558.
12. Broersen, P.M.W., H.E., *On Finite Sample Theory for Autoregressive Model Order Selection.* IEEE Transactions on Signal Processing, 1993. **41**(1): p. 194 -204.
13. Akaike, H., *Fitting autoregressive models for prediction.* Annals of the Institute of Statistical Mathematics, 1969. **21**: p. 243–247.
14. Akaike, H., *A new look at the statistical model identification.* IEEE Trans. Autom. Control, 1974. **19**: p. 716–23.
15. Rissanen, J., *Universal coding, information prediction and estimation.* IEEE Trans. Inf. Theory, 1984. **30**: p. 629–36.
16. Parzen, E., *Multiple time series: determining the order of approximating autoregressive schemes.* Buffalo, NY: Statistical Sciences Division, State University of New York, 1975. **No 23**.