Journal of Sound and Vibration (2001) **240**(5), 909–919
doi:10.1006/jsvi.2000.3290, available online at http://www.idealibrary.com on IDEAL®

A NEW APPROACH TO TIME-DOMAIN VIBRATION CONDITION MONITORING: GEAR TOOTH FATIGUE CRACK DETECTION AND IDENTIFICATION BY THE KOLMOGOROV–SMIRNOV TEST

F. A. Andrade, I. Esat and M. N. M. Badi

Department of Mechanical Engineering, Neural Networks and Dynamic Systems Research Group, Brunel University, Uxbridge. UB8 3PH, England. E-mail: francisco.andrade@brunel.ac.uk

(*Received* 11 *August* 1999, *and in final form* 1 *August* 2000)

This paper introduces a new technique for early identification of spur gear tooth fatigue cracks, namely the Kolmogorov–Smirnov test. This test works on the null hypotheses that the cumulative density function (CDF) of a target distribution is statistically similar to the CDF of a reference distribution. In effect, this is a time-domain signal processing technique that compares two signals, and returns the likelihood that the two signals have the same probability distribution function. Based on this estimate, it is possible to determine whether the two signals are similar or not. Therefore, by comparing a given vibration signature to a number of template signatures (i.e., signatures from known gear conditions) it is possible to state which is the most likely condition of the gear under analysis. It must be emphasised that this is not a moment technique as it uses the whole CDF, instead of sections of the cumulative density function. In this paper, this technique is applied to the specific problem of fatigue crack detection. Here, it is shown that this test not only successfully identifies the presence of the fatigue cracks but also gives an indication related to the advancement of the crack. Furthermore, this technique identifies cracks that are not identified by popular methods based on the statistical moment analysis of the vibration signature. This shows that, despite its simplicity, the Kolmogorov–Smirnov test is an extremely powerful method that effectively classifies different vibration signatures, allowing for its safe use as another condition monitoring technique.

1. INTRODUCTION

In this paper, the problem of fatigue crack identification is tackled. This is a real problem that has already been studied by a number of researchers using a wide range of approaches. Particular attention must be paid to the work of McFadden [1], which uses phase modulations of the gear meshing frequency for crack identification, the work of Boulahbal [2], which uses the wavelet transform and the work of Lin [3], which uses a non-linear dynamical systems approach. All these methods are effective; however, none are simple to implement and interpret.

In the search for a simple but effective technique the Kolmogorov–Smirnov (KS) test was found. This is a time-domain statistical tool suitable for comparing unbinned distributions. It has already been successfully used in fields ranging from astronomy [4] and biology [5] to identification of periodicity in signals [6]. However, the technique has not yet been applied to the analysis of vibration signatures from a condition monitoring point of view.

0022-460X/01/100909 + 11 $35.00/0

This paper shows how the KS test can be used to identify successfully the presence of fatigue cracks on spur gears. In all, four different gear conditions are tested, namely, new gears, normal operating gears, worn-out gears and faulty (fatigue crack) gears. The faulty gears are subdivided into three categories according to fatigue crack advancement.

2. THEORETICAL BACKGROUND

The KS test works on the null hypotheses that the cumulative density function (CDF) of the target distribution, denoted by $F(x)$, is statistically similar to the CDF for a reference distribution, $R(x)$. Hence, it is possible to compare two vibration signatures, and assess if both have the same, or statistically similar, CDFs. Note that the application of this test for condition monitoring assumes that the fault is strong enough to cause a variation in the CDF of the original vibration signature, which is the case of fatigue cracks and many other gear mechanical faults.

The KS test is applicable to unbinned distributions that are functions of a single independent variable. In these cases, the digitised "list of data points" can be readily converted into an unbiased estimator of the CDF, giving some insight about the probability distribution function from which the data were drawn.

It is a fact that different data sets, or different distribution functions, give different CDFs. In this sense, one can establish the likelihood that two sets of data originate from the same distribution function by measuring the differences between their CDFs. A number of statistics can be used to measure the overall difference between two CDFs. In this work, the simple measure of maximum absolute distance between the CDFs was used.

From the two CDFs, $F(x)$ and $R(x)$, a statistic distance D can be calculated. Here, this distance is defined as the maximum absolute difference between $F(x)$ and $R(x)$. Mathematically, this is represented by

$$D = \max_{-\infty < x < \infty} |F(x) - R(x)|. \tag{1}$$

The statistical distance D can be converted into a similarity probability using the KS probability distribution function Q_{ks}. This is defined as

$$prob(D) = Q_{ks}\left(\left[\sqrt{N_e} + 0{\cdot}12 + \frac{0{\cdot}11}{\sqrt{N_e}}\right] D\right), \tag{2}$$

$$N_e = \frac{N_1 N_2}{N_1 + N_2}, \tag{3}$$

where N_e is the effective number of data points and is calculated according to equation (3), N_1 is the number of points in the first data set (reference) and N_2 is the number of points on the second data set (target). The nature of equation (2) is such that it becomes asymptotically accurate as N_e becomes large, but it is also quite good for small values of N_e ($N_e > 5$). In this study, $N_e = 512$ for the numerical example and $N_e = 1024$ for the experimental analysis. Further considerations on the properties and of N_e can be found in reference [7].

Finally, the KS distribution function is defined as

$$Q_{KS}(\lambda) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2\lambda^2}. \tag{4}$$

This is a monotonic function with limiting values of

$$Q_{KS}(\lambda) = \begin{cases} 1 & \text{as } \lambda \to 0, \\ 0 & \text{as } \lambda \to \infty. \end{cases} \tag{5}$$

Hence, if the two vibration signatures are similar (i.e., have statistically similar CDFs), then the similarity probability tends to 1. On the other hand, if the signatures are different, then the similarity probability tends to 0.

3. NUMERICAL EXAMPLE

The numerical examples presented in this section aims to illustrate the processes involved in the KS test. Here, three numerically generated time series are compared. The first two signatures are sine waves (frequency of 200 Hz and amplitude = 1) and the third signature is that of a triangular wave (frequency of 200 Hz and amplitude = 1). Random white noise was added to these signatures. The power of the added white noise is constant for all three signatures and represents 15% of the total signal power. Finally, all signatures were sampled at 1000 Hz, obeying the Nyquist theorem to prevent aliasing errors of the fundamental waveform. However, it must be noted that this sampling frequency will not allow for the detection of the very high frequencies present in the sharp peaks of triangular waves. Figure 1 shows these signatures in the time domain.

As can be seen, at this low sampling frequency, it is impossible to distinguish the three signatures by visual observation of the time domain plots. In fact, the visual observation of these plots suggest that the signatures share a similar basic component and the differences observed can be attributed to noise.

Using the KS test to compare the 200 Hz sine wave with the 200 Hz triangular wave (both sampled at 1000 Hz) gives a similarity probability of 0, indicating that these two signals are in fact statistically different. This similarity probability is found by comparing the maximum absolute difference between the CDFs. In this example $D = 0{\cdot}024$, and is found by plotting the CDF for the two signatures under analysis, and calculating their maximum absolute difference (equation (1)). Figure 2 illustrates this process, showing the CDF for the sine and the triangular wave signal (CDF axis on the left), and the absolute difference '*d*' between the CDFs (*d*-axis on the right). Note that these axis have different scales to allow for the visualisation of the absolute distance between the CDFs.

The statistical distance $D = 0{\cdot}024$ is then fed into equation (2), which gives a probability indicating whether the two CDFs are statistically similar. In equation (2), $N_e = 512$ (i.e., $1024^2/2048$).

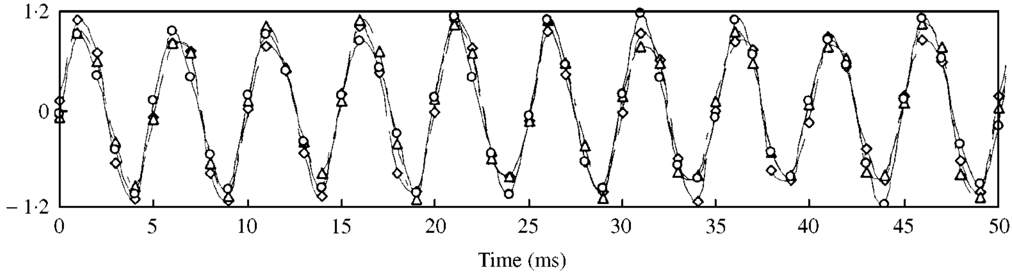

Figure 1. Numerical signatures to be compared: (—◇—), Sin + noise 1; (—△—), Sin + noise 2; (—○—), Tri + noise 3.

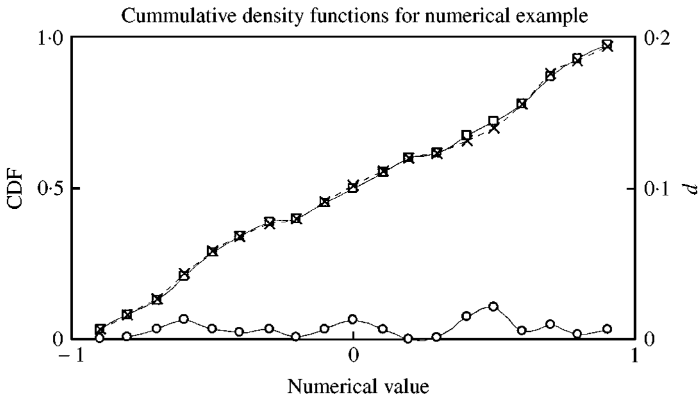

Figure 2. CDF distance for different signatures. (—□—), Sin CDF; (--- × ---), Trian. CDF; (—○—), d.

In this example, the similarity probability between the CDF of the sine and triangular waves equalled zero, showing that the two CDFs are statistically different. Applying the same process to the two sine waves (with added random white noise) a similarity probability of 0·92 was obtained, indicating that the two signals are statistically highly similar, with minor differences due to the added random noise. This example shows the characterization power, which is an important trait of the KS test.

4. EXPERIMENTAL SET-UP AND DATA COLLECTION

In order to investigate the performance of the KS test for fatigue crack identification a test rig was set-up. This modelled a drive-line and allowed for the collection of vibration signatures from the rotating gears. The test rig contains common components present in a real drive-line using meshing gears (i.e., gears, shafts, bearings and couplings). The layout for the rig is shown in Figure 3 and details can be obtained in reference [8].

In this set-up, the driven gear was the object of observation. The same driving gear was used throughout the whole experiment and the driven gear was substituted with test gears in different conditions, namely, new, normal, worn-out and faulty gears (with fatigue cracks of different severity). The gears used in the experiments were manufactured by Davall Gears following the standards DIN3965, and DIN3962. Table 1 summarizes the gear properties.

The gear under observation had a rotational speed of 5 Hz (periodic time 0·2 s) and a constant load of 20 Nm, applied by a pneumatic brake on a brake disc attached to the driven shaft. The vibration signatures were recorded by a magnet-mounted accelerometer placed on the bearing housing adjacent to the driven gear. A sample rate of 5·12 kHz was used to record vibration signatures with 2048 samples. This is equivalent to two full revolutions of the driven gear.

On the driven output shaft, a key and a proximity sensor was used to generate a reference signal (pulse) on the start of each revolution. This signal was also captured by the A/D converter and served as a trigger, indicating the start of each new revolution. This trigger signal was of utmost importance to ensure that all vibration signatures had the same reference start position, allowing for the synchronisation of the vibration signature and shaft position. Figure 4 shows typical vibration signatures of two consecutive revolutions of a normal condition (NO1), and a faulty condition (F3-largest fatigue crack) gear.

The vibration signature blocks (each with 2048 samples or two driven-gear revolutions) were time domain averaged to minimize the effect of noise. The time-averaged signatures

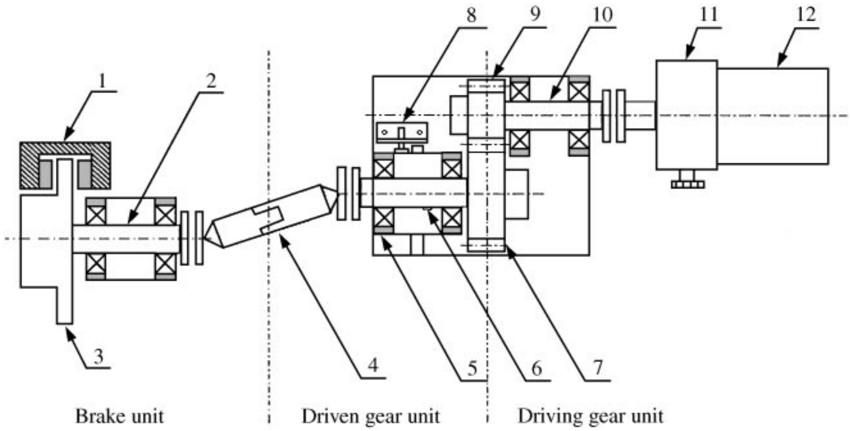

Figure 3. Layout of experimental rig: 1. Brake calliper; 2. Brake shaft; 3. Brake disc; 4. Universal joint; 5. Movable base plate; 6. Output shaft; 7. Driven gear; 8. Spacer block; 9. Driving gear; 10. Input shaft; 11. Kopp variator®; and 12. AC motor.

Table 1

Characteristics of test gears

Parameter	Driving gear	Driven gear
Type	MA25-20S	MA25-32S
Number of teeth	20	32
Module	2·5	2·5
Face width (mm)	25	25
Pressure angle (deg)	20	20
Helix angle (deg)	0	0
Pitch diameter (mm)	50	80
Material (mild steel)	EN8	EN8

were then processed using the KS test and existing statistical measures (moment analysis). Figure 5 compares the averaged signal and the raw vibration signature for a good-gear condition.

The faulty gears were obtained by introducing fatigue cracks to normal operating driven gears. For this a disc cutter of diameter 55 mm and thickness of 0·3 mm was used. The cut aimed at replicating a crack at the critical tooth section that grows along the critical stress line [8]. Figure 6 shows how the three different cuts simulate the crack advancement.

Note that the cuts F1 and F2 stop at 8 and 16 mm (respectively) across the tooth face. This is indicated in the diagram on the right section of Figure 6. For the large fatigue crack (F3) the cut covers the whole face of the tooth (25 mm). For a full description of the cut geometry and dimensions refer Table 2. Figure 7 illustrates the faults being investigated.

In all, vibration signatures (each consisting of 2048 samples) from seven driven gears were collected. The conditions of these gears were as follows:

- One brand new gear (BN),
- Two gears in normal operating condition (NO1, NO2),
- One worn-out gear (WO—showing signs of pitting and scoring),
- Three gears with introduced fatigue cracks of different severity (F1, F2 and F3).

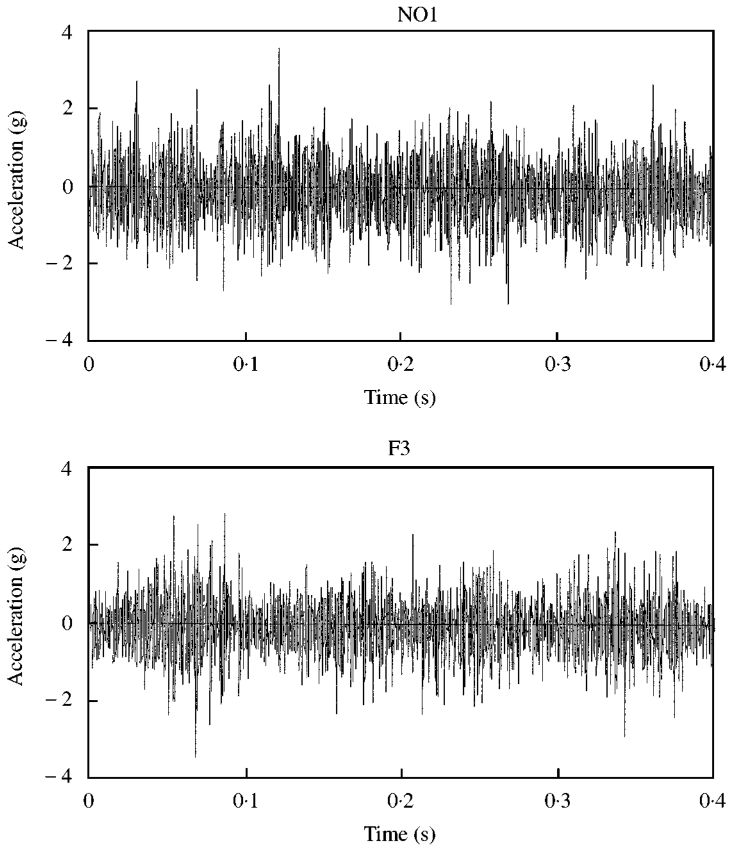

Figure 4. Typical vibration signature for a good and a faulty condition gear.

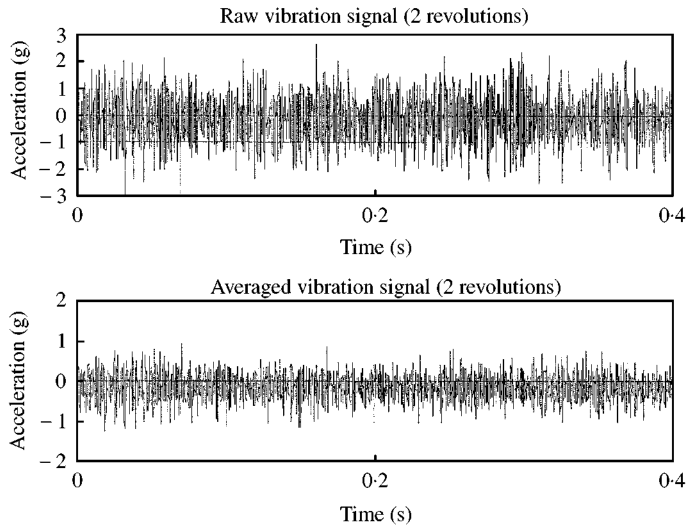

Figure 5. Typical vibration signatures: raw and time-averaged signal.

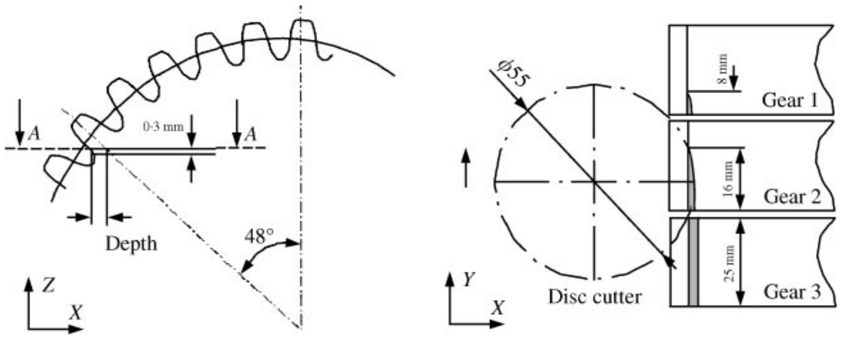

Figure 6. Schematic diagram of fatigue cut on spur gears.

TABLE 2

Cut geometry and illustration of crack angle

Gears	Depth (mm)	Width (mm)	Thickness (mm)	Angle (deg)
F1	0·8	8	0·3	40
F2	1·6	16	0·3	40
F3	2·4	25	0·3	40

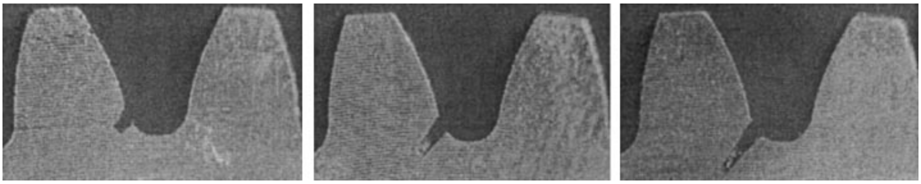

Figure 7. Faults under investigation.

It must be noted that two gears in normal operating condition were used to ensure that any changes in the KS value were due to the gear condition, and not simply due to the different gears, as the comparison of two similar normal condition gears should lead to a high similarity probability.

From the experimental rig described, 48 vibration signatures (each with two driven gear revolutions) were recorded for each gear condition. The signatures were locked and aligned onto the output shaft rotation by means of the trigger signal from the proximity sensor. The "aligned" signatures were grouped into fours, for the time domain average. Hence, for each gear condition 12 time-averaged vibration signatures were obtained. To these signatures the KS test was applied. Also, tests using existing techniques such as kurtosis, skewness, form- and crest-factor were performed, to assess the effectiveness of the KS test against other existing techniques.

Finally, it must be noted that four averages do not fully separate the vibration of the driven gear from that of the driving gear; for this, many more cycles would have to be included in the time average. This conscious choice is aimed at showing the effectiveness of the performance of the KS test in systems with little past information (vibration signatures) available.

5. EXPERIMENTAL RESULTS

Figure 8 shows the results obtained from the calculation of standard statistical measures from 12 vibration signatures (from each gear condition). Note that each signature includes two-gear revolutions (i.e., 2048 samples).

As can be seen, none of the statistical moment measures reliably identify the different gear conditions. Also, these measures are not able to classify the severity of the different faults, nor classify the worn out gear (WO) and the faulty gears (F1, F2 and F3).

Figures 9 and 10 show the results obtained by using the KS test to compare the vibration signatures. Firstly, Figure 9 uses a good gear condition signature as the reference signal for the comparison. Secondly, Figure 10 uses a vibration signature from the gear condition F3 as a reference signal (crack with greatest severity). The results are displayed in the form of a similarity probability, where 1 indicates high similarity (hence signals come from gear with the same/similar condition) and 0 indicates low similarity.

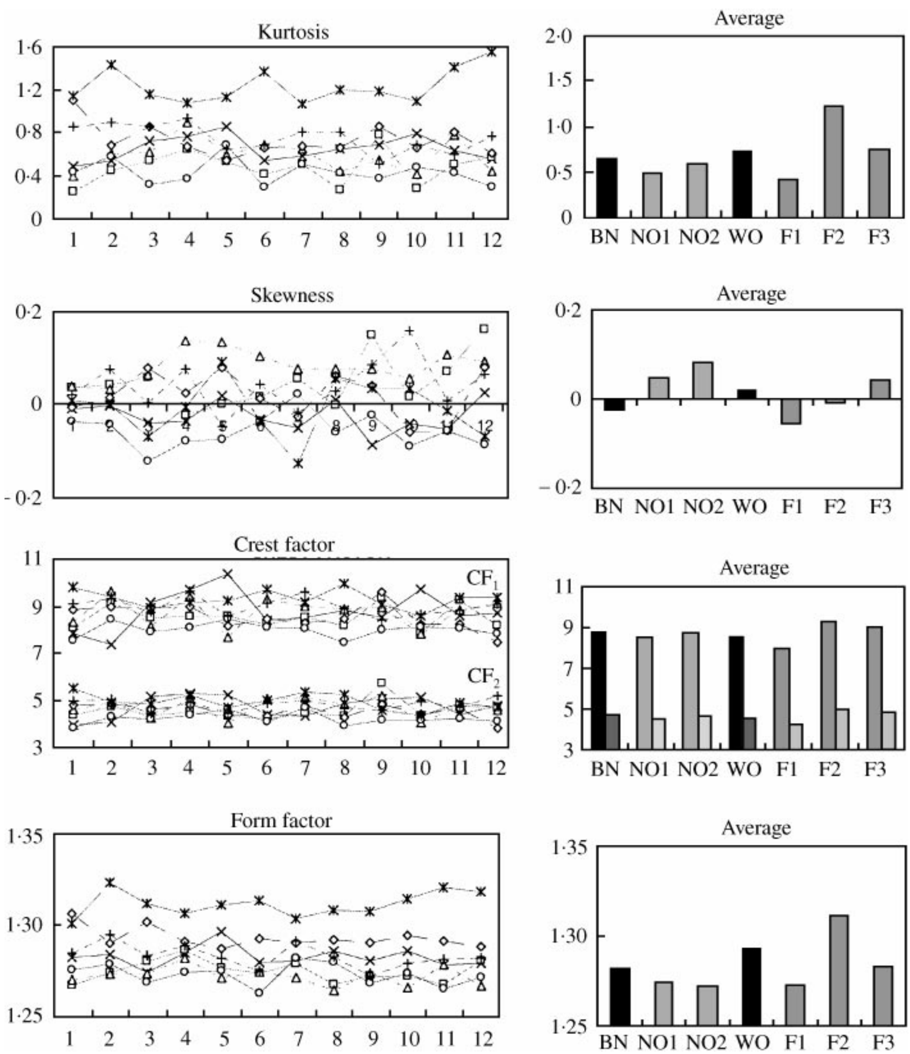

Figure 8. Standard Statistical measures of experimental data: (—×—), BN; (—□—), NO1; (—△—), NO2; (—◇—), WO; (—○—), F1; (····· × ·····), F2; (- - + - -), F3.

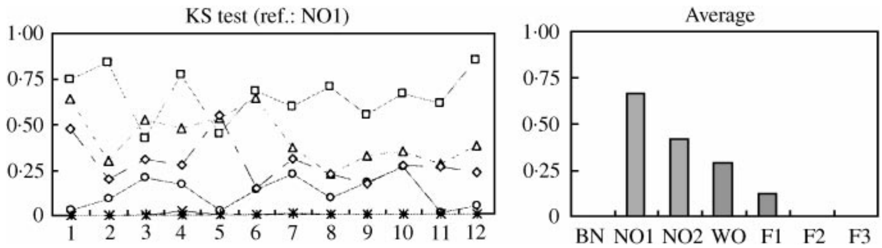

Figure 9. KS test using NO1 as a reference signal: (—×—), BN; (—□—), NO1; (—△—), NO2; (—○—), F1; (····· × ·····), F2; (- - + - -), F3; (—◇—), WO.

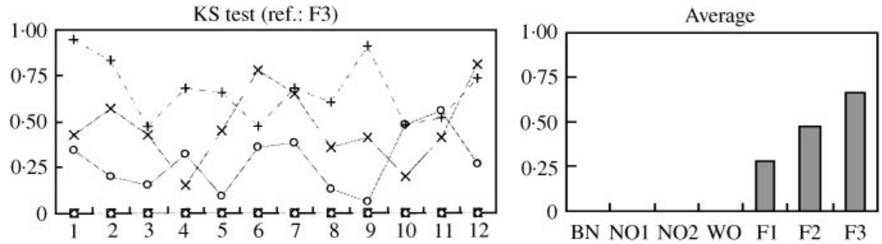

Figure 10. KS test using F3 as a reference signal: (—×—), BN; (—□—), NO1; (—△—), NO2; (—○—), F1; (····· × ·····), F2; (- - + - -), F3; (—◇—), WO.

TABLE 3

Experimental results

Reference signal	Comparison						
	BN	**NO1**	NO2	WO	F1	F2	**F3**
NO1	0·00	**0·63**	0·38	0·26	0·10	0·00	0·00
F3	0·00	0·00	0·00	0·00	0·29	0·53	**0·64**

It must be noted that the reference signal used is not included as a test signal. Hence, a similarity of 1 is not expected, as this would indicate that the signals are statistically identical. The collected vibration signatures will contain random noise, which will prevent signals from being identical.

Finally, based on the results from the analysis of these 12 signatures, the same tests (i.e., first with NO1 and secondly with F3 as a reference signal) were performed with all the vibrations signatures. Hence, 24 vibration signatures (firstly from NO1 and then from F3) were used for reference and 24 signatures from all other gear conditions were used as test signals. An average of the results is included in Table 3.

As can be seen when using NO1 as a reference signal the highest similarity arises when using other cycles from the same gear condition (NO1 and NO2) as a test signal. Similarly, when using the faulty gear condition (F3) as a reference signal, the highest similarities arise from the faulty condition gears. Once again, it is important to note that even similar gear conditions do not give a similarity of 1. This is attributed to the noise present in all the signatures.

6. DISCUSSION

This paper showed how the KS test is able to distinguish between vibration signatures originating from gears in different conditions. The KS test, performed more robustly than other existing time-domain techniques based on the calculation of the vibration signature statistical moments.

However, like the moment techniques, the KS test is strongly dependent on the gear loading and rotational speed, as variations in these parameters affect greatly the vibration amplitude and the fundamental frequency (meshing frequency) in the vibration signatures of the gears under observation. Preliminary tests suggest that the dependence on load can be minimized by normalizing the vibration signature amplitude. The results obtained from these tests are not conclusive, and further tests must be carried out to confirm this hypothesis.

It must be noted that the experiment performed here used a simple drive-line consisting of only one pair of gears. In a more complex drive-line it is possible that the transients induced by the gear fault may be masked by the vibration of other undamaged gears. This drawback is intrinsic to any digital signal processing technique applied to vibration condition monitoring. Ways around this problem include studying the effectiveness of the placing sensor on the gearbox casing, and also the use of time-domain averaging to enhance the vibration from any specific gear under analysis.

Finally, a corollary of the results presented here, suggests the extension of this technique to a comparison of time–frequency and/or time-scale maps. A generalization of the KS test to a two-dimensional distribution has already been suggested by Fasano [9], based on earlier research by Peacock [10]. Again this would form part of a larger automated fault identification system. The main advantage of this method over the usage of neural networks for automated pattern recognition lies in the fact that this technique does not require the computational time and expense intrinsic to the training process of neural networks.

7. CONCLUSIONS

It has been shown that the KS test effectively distinguishes vibration signatures by comparing their CDFs, and can be used for the identification of tooth fatigue cracks on gears. Furthermore, if a template of possible gear conditions is available, it can be used to estimate fault advancement. This is shown by the clear trend obtained when using the vibration signatures from a gear with the largest fatigue crack (F3) as a reference signal. In this test, a trend showing higher similarity probability for more severe cracks is clearly seen.

The results obtained here also suggest that if a number of vibration signatures from different gear conditions are available, then it is possible to classify different types of faults, leading to automated fault detection systems.

ACKNOWLEDGMENTS

This research was partially supported by the Brazilian Federal Government Agency, CAPES.

REFERENCES

1. P. D. McFadden 1986 *Transactions of the American Society of Mechanical Engineers, Journal of Vibration, Stress and reliability in Design* **180**, 165–170. Detecting fatigue cracks in gears by amplitude and phase modulations of the meshing vibration.

2. D. Boulahbal, M. F. Golnaraghi and F. Ismail 1997 *Proceedings of DETC'97*, *American Society of Mechanical Engineers Design Engineering Technical Conference Gear crack detection with the Wavelet Transform.*
3. D. C. Lin, M. F. Golnataghi and F. Ismail 1997 *Journal of Sound and Vibration* **208**, 664–670. The dimension of the gearbox signal.
4. B. X. Santiago, R. A. W. Elson and G. F. Gilmore 1996 *Monthly Notice of the Royal Astronomical Society* **281**, 1363–1374. HST photometry of 47 Tuc and analysis of stellar luminosity function in Milky Way clusters.
5. M. Akay, Y. M. Akay, P. Cheng and H. H. Szeto 1995 *Biological Cybernetics* **72**, 431–437. Investigating the effects of opioid drugs on electrocordical activity using wavelet transform.
6. H. Arsham 1997 *Stochastic Hydrology and Hydraulics* **11**, 323–330. A test sensitive to extreme hidden periodicities.
7. R. Von Mises 1964 *Mathematical Theory of Probability and Statistics*. New York: Academic Press, Chapters IX(C) and IX(E).
8. S. Engin 1998 Ph.D. Thesis, University of Hertfordshire, England. Condition monitoring of rotating machinery using wavelets as a pre-processor to artificial neural networks.
9. G. Fasano and A. Franceschini 1987 *Monthly Notices of the Royal Astronomical Society* **225**, 155–170. A multidimensional version of the Kolmogorov–Smirnov test.
10. J. A. Peacock 1983 *Monthly Notices of the Royal Astronomical Society* **202**, 615–627. Two-dimensional goodness-of-fit testing in astronomy.