# The Application of Fourier Analysis to Forecasting the Inbound Call Time Series of a Call Centre

**Bruce G. Lewis[a], Ric D. Herbert[b] and Rod D. Bell[c]**

[a]NSW Police Assistance Line, Tuggerah, NSW 2259, e-mail:lewi1bru@police.nsw.gov.au, Faculty of Science and Information Technology, University of Newcastle, Ourimbah, NSW 2252

[b,c]Faculty of Science and Information Technology, University of Newcastle, Ourimbah, NSW 2252

**Abstract:** The New South Wales Police Assistance Line (PAL) is a 24-hour inbound telephone call centre available to police and the community. One of the challenges in running such a centre is the scheduling of staff and resources to meet the incoming call demands. The call arrival process of an inbound call or contact centre is a time series that contains seasonal patterns, cyclic patterns and trends. We use fast Fourier transforms (FFTs), a powerful signal processing technique, for the analysis of time series data. Controlled examples are used to assess the utility of the process which is subsequently applied to the PAL time series call incoming data. This has identified a number of underlying call patterns which will be used to assist the scheduling of staffing resources.

***Keywords:*** *Call Centre, Fourier analysis, Fast Fourier Transform, NSW Police Assistance Line.*

## 1. INTRODUCTION

The New South Wales (NSW) Police Assistance Line (PAL) is a 24-hour inbound telephone call centre available to police and the community. PAL operates as a single virtual call centre across the state of NSW from two geographically differentiated sites at Tuggerah on the NSW Central Coast and Lithgow in the Central West of NSW. PAL is a niche facility in the call centre industry, being used for the reporting of non-urgent crime and incidents, for providing police-related information to the community, report crime and for providing an intelligence source for the NSW Police. In the ensuing years it is planned to move the facility to an inbound multimedia contact centre to take calls from its customers via different media including telephone, facsimile, Internet or email. This may change the profile of the incoming traffic and it will be necessary to understand the underlying incoming call patterns and their exogenous drivers to ensure that sufficient staff and facilities are available to maintain the current high standards (Lewis *et al.*, 2002). To achieve this necessitates an understanding of the call arrival time series.

One method, frequency analysis, is used in communications and Engineering and its potential will be investigated for this application. Recent literature relating to call centre forecasting methods (Hughes, 1995; Klungle and Maluchnik, 1997-'98; Bianchi *et al.*, 1993) revealed that no work has been undertaken in using Fourier analysis to develop a process for forecasting a call centre's incoming calls time series. However, since Young *et al.* (1999) state that, periodic, seasonal or cyclical effects are exhibited in many socio-economic and environmental time series and Cleveland and Mayben (1997) state that virtually all inbound call centres, including emergency services centres, have distinct calling patterns which can usually be detected down to at least a half-hour. Fourier analysis may reveal the cyclic and seasonal variations from these patterns even though the precision with which calls will arrive cannot be predicted.

This paper develops a model in which the components of a time series are determined in the frequency domain by using the Fast Fourier Transform (FFT) as the enabling tool. This development has been undertaken using Matlab.

## 2. TIME SERIES ANALYSIS

It is well known that forecasts can be developed using a past history of data or time series consisting of one or more of a long term trend, a seasonal variation, a cyclic variation and random effects. The seasonal variation is often tied to annual cycles and the presence of random components prevents the exact prediction of future values (see for example the text Lawrence and Pasternak, (1998)).

Klungle and Maluchnik (1997-'98) have found that for call centre forecasting, the regression model provides the flexibility to deal with causal variables and predicts the overall trend, systemic changes and spikes with the ability to combine time series analysis and causal being the model's main advantage. However, they found that the practical application of regression for medium to long term forecasting is difficult.

This motivates us to investigate the use of FFTs to predict PAL inbound calls and in this paper we present some preliminary results.

## 3. FOURIER ANALYSIS

In 1822, Joseph Fourier completed his work on the Théorie Analytique de la Chaleur (The Analytical Theory of Heat) in which he introduced the series

$$y = \frac{1}{2}a_0 + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \ldots \quad (1)$$

as a solution (D.J.S, 2001). This process was to later become known as Fourier analysis in which it is stated that any periodic function can be expressed in the form (1) (Betts, 1972).

An improvement, the Fast Fourier Transform (FFT) provides significant reductions in computation time (Stremler, 1990). In his paper, Gonzalez (1990) explains the practical aspects of implementing the FFT. The output of the FFT process is a series of equally spaced frequency lines where there are half as many frequency lines as data points. This is because the FFT generates complex numbers and each line contains two pieces of information, being amplitude or phase, or real and imaginary data.

In undertaking the FFT process, Gonzalez (1990) states sampling at a consistent time interval is important to prevent distortion and warns against aliasing and leakage. Furthermore, unless data is sampled at a rate that is at least twice as fast as the maximum frequency under observation, unwanted false low frequency or alias components will occur. This may be overcome by applying an anti-aliasing filter to the data at one half of the sampling frequency.

Where the sampled time history is not an exact multiple of the sample period, leakage occurs which causes unwanted components or smearing around the wanted frequency components. This occurs if the data being analysed does not contain an integral number of cycles of the input wave. A technique called windowing can be used to reduce this problem by multiplying the time record with a window function to concentrate on the middle of the time record and ignore the ends (Gonzalez, 1990).

## 4. THE MODEL

The model used to predict call arrivals, $y(t)$, is

$$y(t) = a_1 + a_2 t + \sum_{i=1}^{n} b_i \sin(2\pi f_i t + \theta_i) \quad (2)$$

where $a_1$, $a_2$, $b_i$, $f_i$ and $\theta_i$ are the parameters to be estimated.

For a given time series, the frequency domain is used to determine the gradient, $a_2$, and vertical displacement, $a_1$, from the abscissa, of any trend line present.

The estimation program was written using Matlab. Using a Fast Fourier Transform algorithm, the incoming time series is loaded and sampled from a specific start time and for a specific length. The FFT of the sample is then taken and a number of non-zero frequencies, selected by the user, are extracted. The remaining frequency spectrum is then used to estimate the trend line and abscissa. This is done by a non-linear fitting procedure in the frequency domain.

Predicting a future length of the series is undertaken from a sample number selected by the user. All calculations are referenced from the beginning of the data sample time series.

For each of the selected frequency components, the phase was determined at the start of the prediction period and a time series generated. The components were then summed for the total series. The algorithm used to calculate the predicted time series is shown in (2).

$\theta_i$ is calculated from

$$\theta_i = (2\pi \frac{f_i}{f_s})(n_{ss} - 1) \quad (3)$$

where $f_s$ is the sample frequency, $n_{ss}$ being the start sample number.

## 5. RESULTS

Before using the PAL data we present some results form a controlled experiment.

## 5.1. Control data

The control data comprised a trend line with a gradient of 1 and a displacement of 1.5, and three sine waves of frequencies 2Hz, 5Hz and 7Hz with respective amplitudes of 2, 1 and 1.7. These were sampled at a rate of 128Hz with a sample size f 256 starting at sample 77 from the start of the series. In each of the following figures, the dotted line shows the original time series data. Samples and predictions are shown as continuous lines.

Figure 1 shows the input time series and the sample. Figure 2 is the FFT of the time series sample in which is seen the three components of the control signal and the trend line.

Initially the dominant frequencies are temporarily removed by selecting the largest non-zero components as determined by the user. The zero frequency component is retained since it holds valuable information about the gradient and abscissa displacement of the trend line.

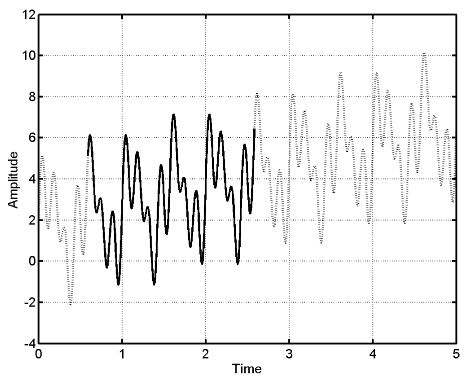


**Figure 1**. Time series and data sample

The trend line (gradient and displacement) is estimated by fitting the reduced frequency domain data. The frequency spectrum of the fitted trend line is then algebraically subtracted from the band limited data sample. Figure 3 shows the original spectrum with the trend spectrum removed.

Using this information and knowing at which point the sample period commenced (equation (3)), equation (2) is used to predict the time series at the end of the sample period (Figure 4). In this case, the prediction and original data are the same since the parameter estimation procedure has resulted in values equal to those used to generate the input data.

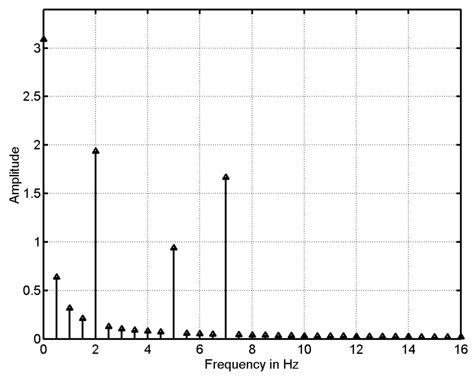


**Figure 2.** Sample Fast Fourier Transform

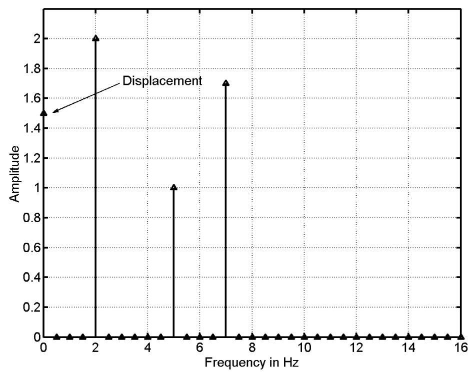


**Figure 3.** Filtered Fast Fourier Transform of Band limited data without trend

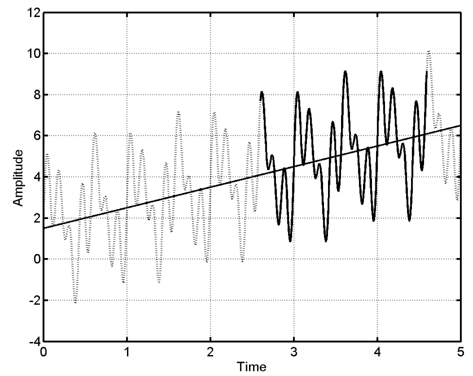


**Figure 4.** Predicted sample continuation

## 5.2. PAL data

The PAL data set comprised one year of data from 1 July 2001 to 30 June 2002 and is shown in Figure 5. In this figure cyclical patterns can be seen. Figure 6 and Figure 7 show a finer grain of the data. Figure 6 shows the call pattern for August 2001 in which a seven-day cycle is clearly seen.

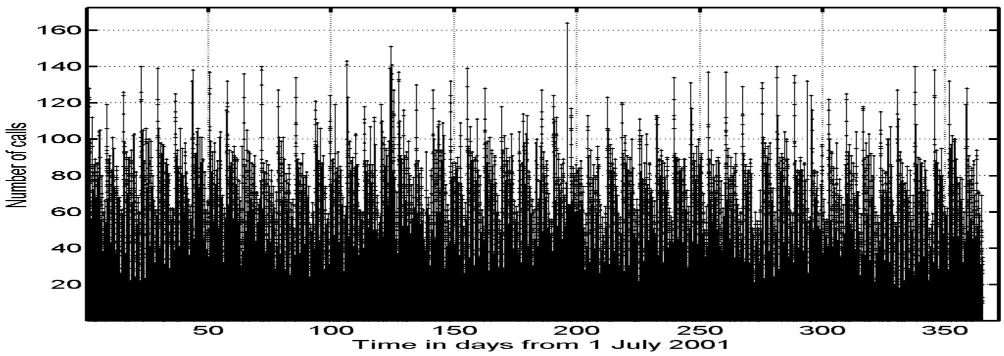


**Figure 5. Call data for year 2001-2002**

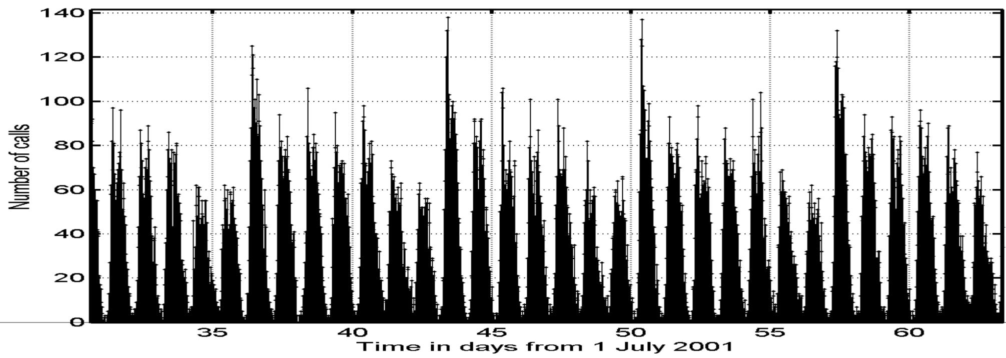


**Figure 6 Call pattern for August 2001**

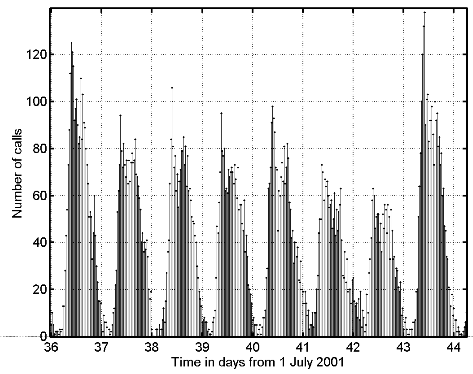


**Figure 7 Typical weekly call pattern**

Figure 7 shows the cyclical nature of a weekly call pattern for the second week in August 2001. The daily pattern can be seen in this figure.

The FFT was evaluated at a sample frequency of 48 samples per day. Due to the complexity of the time series, the largest 11 frequencies were selected for evaluation. Figure 8 shows the spectrum with the dominant components indicated for 52-weeks of data. From the figure the strong daily pattern can be seen. Not so obvious from the time series data is the 7-day, 12-hour, 8-hour and 6-hour cycles. Further analysis of the data for data sample lengths of one-week, four-weeks, 12-weeks, 26-weeks and 52-weeks has been made. Table 1 shows the resulting frequency components together with their associated amplitudes, which correspond to the peak number of calls. This correlates with the workload experienced by the call centre where there is a known daily cycle, where more calls are received in the daylight hours, and there is a weekly cycle with Mondays being the busiest day. It was interesting to note that there was no significant component for seasonal cycles or an annual cycle which would have corresponded to a frequency of 0.0027 cycles per day. The results show that, in the long run, there was no increase in the call rate over the twelve months and that the average displacement of 40 calls compared well with the average of 39 calculated from the raw data. These results are encouraging and more work is required to understand the nature of frequencies with respect to the PAL business.

To test the model's predictive power, a start time of day 231 at 11:30 pm, which corresponds to 16 February 2002, was selected arbitrarily. Figure 9 shows the result where the histogram is the call centre data and the continuous line is the prediction. Figure 10 shows the prediction over a longer time period.

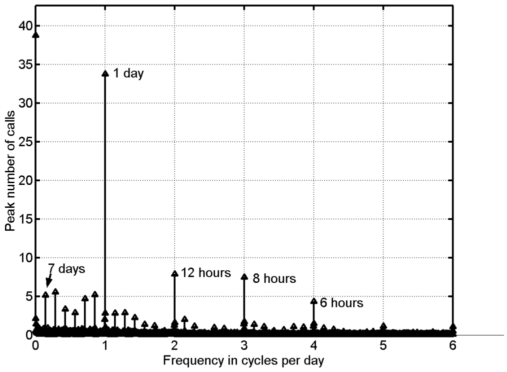


**Figure 8 Spectrum of data for year 2001-2002**

From the figures, the model predicts the future call arrival pattern quite accurately. The accuracy is more than sufficient to aid in the scheduling of staff and resources for the call centre.

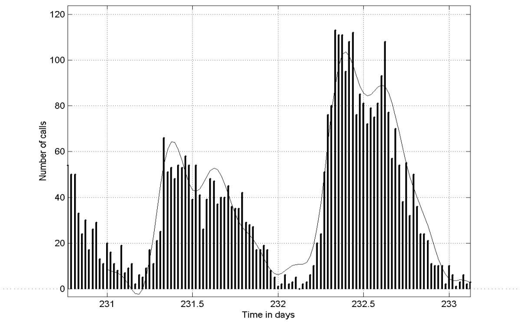


**Figure 9 Predicted and actual calls**

## 6. CONCLUSIONS

Working in the frequency domain overcomes many difficulties encountered in the time domain. Individual frequencies, blocks of frequencies or groups of non-related frequencies can be readily manipulated.

This paper has shown that for periodic data with a trend, the technique of fitting trend data in the frequency domain using FFTs has been very effective at differentiating between cyclical data and trends.

The results using PAL data are encouraging and have demonstrated that the technique has potential for forecasting call volume. In particular, cyclical patterns that were not obvious in the time domain data have become very clear in the frequency domain.

Further research will be undertaken in the areas of performance under narrow band and wide band conditions, identifying, and correlating the time series components against business cycles and natural phenomena, sample size and sample frequency dependency, non-linear trends, short- versus long term forecasting accuracy and, examining different methods of assessing trend and abscissa displacement.

## 7. ACKNOWLEDGEMENTS


We wish to acknowledge the support of the NSW Police and the Director of the NSW Police Assistance Line in supporting the research and the provision of the data used for analysis. We also acknowledge the helpful comments from our reviewers.


## 8. REFERENCES


Betts, J. A. *Signal Processing, Modulation and Noise*. The English Universities Press Limited. London, 206, 1972.

Bianchi, L., J. E. Jarrett and R. C. Hanumara, Forecasting incoming calls to telemarketing centers. *The journal of Business Forecasting*, 3–11, 1993.

Cleveland, Brad and Julia Mayben, Call Center Management on Fast Forward: Succeeding in Today's Dynamic Inbound Environment. Call Center Press. Maryland, USA, 1997.

D.J.S, Encyclopaedia Britannica: Fourier, (Jean-Baptiste-),Joseph, Baron, Oxford University Press, 1994-2002.

Gonzalez, Ozzie, Pitfalls of signal analysis. *Mechanical Engineering* pp. 72–74, 1990.

Hughes, C. Four steps for accurate call-centre staffing. *HRMagazine,* 87–89, 1995.

Klungle, R. and J. Maluchnik, Call center forecasting at AAA Michigan. *The Journal of Business Forecasting,* 8–13, 1997-'98.

Lawrence, J. A. and B. A. Pasternak, Applied Management Science: A Computer Integrated Approach for Decision Making. John Wiley & Sons, Inc. U.S.A. 386-412, 1998.

Lewis, B.G., R.D. Herbert and W.J. Chivers, The migration of a call centre to a contact centre, paper presented at the First International Conference on Information Technology & Applications (ICITA 2002), Bathurst, Australia, 25-28 November 2002.

Stremler, F. G., Introduction to Communications Systems: Third Edition, Addison-Wesley, USA, 1990.

Young, P.C., D.J. Pedregal and W. Tych, Dynamic Harmonic Regression. *Journal of Forecasting.* UK, 1999

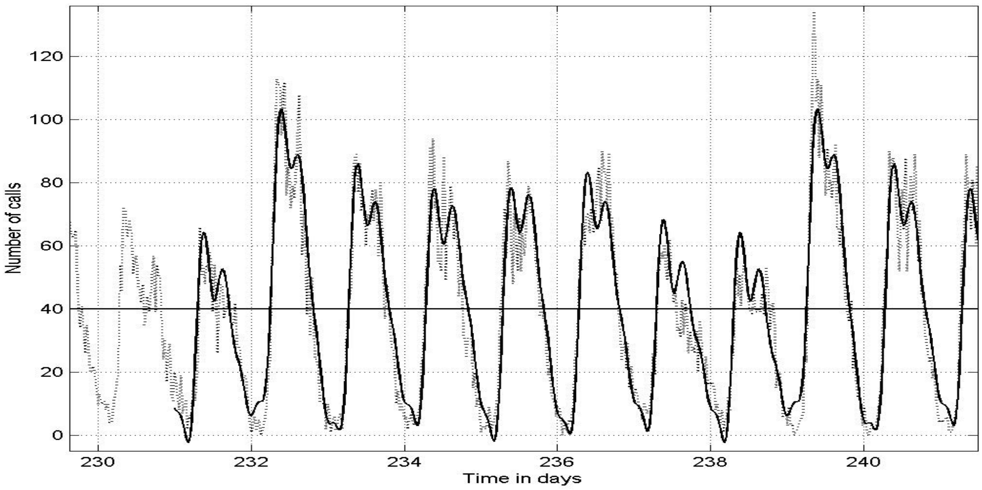


**Figure 10 Predicted calls (solid line) superimposed on the call data (dotted line)**

**Table 1 Dominant frequency components and peak amplitudes for data of varying weekly lengths**

| Frequency | | Peak number of calls | | | | | Average |
|---|---|---|---|---|---|---|---|
| | | Number of weeks analysed | | | | | |
| Cycles per day | Period | 1 | 4 | 12 | 26 | 52 | |
| 0.1429 | 7 days | 5.1240 | 5.4112 | 4.8743 | 4.9664 | 5.1410 | 5.1034 |
| 0.2857 | 3.5 days | 7.4578 | 6.0107 | 5.9940 | 5.8474 | 5.5372 | 6.1694 |
| 0.4268 | 2.34 days | 3.2622 | 3.6317 | 3.8040 | 3.6796 | 3.3336 | 3.5422 |
| 0.5714 | 1.75 days | 3.6216 | 3.1574 | 3.0428 | 3.1116 | 2.8711 | 3.1609 |
| 0.7143 | 1.4days | 5.0460 | 4.8893 | 5.0155 | 4.7937 | 4.6575 | 4.8804 |
| 0.8571 | 1.17 day | 5.5673 | 5.5265 | 5.3337 | 4.9470 | 5.2096 | 5.3168 |
| 1.0000 | 1 day | 35.8447 | 35.9452 | 35.3448 | 34.5606 | 33.7915 | 35.0974 |
| 1.2857 | 18.7 hrs | 4.0344 | 3.6146 | 3.2183 | 2.9717 | 2.8553 | 3.3389 |
| 2.0000 | 12 hrs | 8.2790 | 8.3554 | 8.2848 | 8.2178 | 7.8762 | 8.2026 |
| 3.0000 | 8 hrs | 8.5798 | 8.2040 | 8.2774 | 7.7365 | 7.4712 | 8.0538 |
| 4.0000 | 6 hrs | 5.6268 | 5.4034 | 5.6676 | 4.6838 | 4.3530 | 5.1469 |
| | Trend gradient | 0.3400 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0680 |
| | Displacement | 39.8016 | 40.8609 | 40.1138 | 40.3645 | 38.7758 | 39.9833 |