

Assessment of the Effectiveness the Use of Spectral Analysis in Forecasting Of Sales Quantity of Hard Coal.

Marcin Popczyk, Dr. Eng.*, Anna Manowska, Dr. Eng.*, Rafał Jendruś, Dr. Eng.***

*Silesian University of Technology, Faculty of Mining and Geology, Akademicka 2A, 44-100 Gliwice, Poland, www.polsl.pl

**Geotechnical Services Company „GeoRock”, Armii Krajowej 14/1/10, 41-943 Piekary Slaskie, Poland, www.georock.pl

Abstract: Coal will be a cheaper alternative than natural gas, and oil to generate electricity in many regions, but in the long term essential for determining the future will accepted strategies to improve efficiency, reduce air pollution at the local level and mitigating the effects of climate change. Currently, the Ministry of Economy is preparing an update "Polish Energy Policy until 2050". The new document will consider the most important elements in the energy sector such as min. use of gas from unconventional sources, climate challenges, the construction of a common energy market in the European Union or the use of European funds in the new financial perspective 2014-2020. Sales forecasting coal is the subject of this paper, is a problem whose solution should look for today. This paper presents a prediction method that can be used to solve specific problems of forecasting sales volumes of coal, which is the result of logical conclusions following areas: mining, predictive theory, energy security and social security.

I. Introduction

Poland is not only one of Europe's traditional hard coal producers; it was once one of the main suppliers to the world hard coal market. The country had assumed the leading role among European coal mining countries in 1972 with an output of 150.7 Mt and, until 1979, was the world's second largest coal exporter after the USA, selling 41.4 Mt in that year. Although its role as an exporting country was already fading in the 1980s, the extraction was maintained at a significant level (1988: 193 Mt) compared with other European countries. It was not until the political turnaround in eastern bloc countries associated with a growing market-economy that Poland experienced, in the early 1990s, a contraction of its hard coal mining industry, a process that had already begun in Western Europe 20 years previously. Thus, the extraction amounted to a mere 78,1 Mt in 2012 and has stabilized in recent years. Poland's coal is currently in a better competitive situation than in the past thanks to high world market prices. 93 % of the country's power generation is based on hard coals and lignite, three fifths being accounted for by hard coal and two fifths by lignite[2, s.91].

Total extraction was increased by 2.8 Mt in 2012 to some 78,1 Mt. The scale-down in production came mainly at the expense of seaborne exports, while domestic sales remained largely stable in 2012. The coking coal group Jastrzebska, by contrast, was able to increase its output and exports. Poland's mining sector urgently needs investment to maintain extraction. By 2010, the Polish state proposed to discontinue its subsidies, which were some USD 2.0 – 2.3/t for steam coal. The government's concept for 2007 – 2013 was set to make Poland's pits competitive and subsidy-free. The extraction was expected to fall to 77 - 78 Mt in 2010 and to 70 Mt in 2020 [2, s.92]. Poland's mining sector faces further change. Although output has stabilized in recent years, further falls must be expected since too little is being invested in the development of new reserves. Rising wages without matching advances in productivity, are also making the situation more difficult. For its exports to be competitive, Poland needs competitive prices. Increasingly, the country is confining itself to neighboring countries as buyers, i.e. the Baltic states, Germany and Austria [2, s.10].

The paper presents a method of forecasting sales of coal using harmonic analysis, as it becomes necessary to determine the capacity of the mining industry, which will satisfy the changing demand for coal. This issue is important not only in the analysis and evaluation of the country's energy security but also in the increase of coal exports.

II. Method of spectral analysis

To describe the concept of the method of spectral analysis, there should be introduced the term: function of the power spectrum, which is the first derivative of the spectrum distribution $F(\omega)$ of the stochastic process. This function can be expressed as[1, s.101]:

$$dF(\omega) = f_2(\omega) + \square_i^2 \quad [1]$$

$$f_2(\omega) = dF_2(\omega)/d(\omega) \quad [2]$$

where $f_2(\omega)$ is the spectral density of the process $X_2(t)$ that 'hides' the cyclic component and σ_r^2 is the r-th variance of the process $X_1(t)$ of strict periodicity.

The relationship (3.34) can be used as a tool for an analysis of periodicity for two main reasons:

- (a) if the cyclic component equals zero ($X_1(t)=0$), then the function $dF(\omega) = f_2(\omega)$, which means it covers the spectral density of the process $X_2(t)$, which is an absolutely continuous function
- (b) if there is a cyclic component ($X_1(t)\neq 0$), then the function $dF(\omega)$ is not absolutely continuous; in points $\omega = \omega_r$ the value of the function jumps up because to the value of the function $f_2(\omega)$ is added the variance σ_r^2 of r-th component of the process $X_1(t)$.

In practice, when a graphical picture of the function $dF(\omega)$ is drawn, it is easy to notice such points ($\omega_r; r = 1, 2, \dots, s$) where the value of the function increases drastically. Therefore, one can say on the periodicity of the process in its point is $2\pi/\omega_r$. A difficult problem arises when in some points the function increases only slightly. Unfortunately, estimation the function $dF(\omega)$ is usually done with a certain accuracy only and there is no clear indication whether the observed increment of the value of the function is significant or not. However, there are several methods that can be used to dispel any doubts. Different authors recommend different tests; however, many of them are complicated procedures.

1. The idea of harmonic analysis

If the observed time series $y(t_i); i = 1, 2, \dots, n$ has no trend, it can be expanded in a Fourier series. In mathematics, a Fourier series decomposes a periodic function into the sum of simple oscillating functions, namely sines and cosines. Following this line of reasoning, we can write [1, s.101]:

$$y(t_i) = \frac{1}{2} a_0 + \sum_{i=1}^n (a_i \cos \omega_i t_i + b_i \sin \omega_i t_i) \quad [3]$$

where a_0, a_i, b_i are the Euler-Fourier coefficients.

The estimators of these coefficients are as follows:

- The expected value estimator [1, s.102]

$$\hat{a}_0 = \frac{1}{n} \sum_{i=1}^n y(t_i) \quad [4]$$

- Further estimators of the Euler-Fourier coefficients [1, s.102]

$$\hat{a}_j = \frac{2}{n} \sum_{i=1}^n y(t_i) \cos \frac{2\pi i j}{n} \quad [5]$$

$$\hat{b}_j = \frac{2}{n} \sum_{i=1}^n y(t_i) \sin \frac{2\pi i j}{n} \quad [6]$$

for $j = 2, \dots, n/2$.

The values of these statistics can be applied to verify the hypothesis that a cyclic component is significant in a given sample.

Consider the square amplitude of the process. It is given by the equation:

$$\hat{A}_j^2 = \hat{a}_j^2 + \hat{b}_j^2 \quad [7]$$

Its expected value is

$$E(\hat{A}_j^2) = A_E^2 = 4 \frac{\sigma^2}{n} \quad [8]$$

where σ^2 is the variance of the process.

By replacing the unknown variance σ^2 with its unbiased estimator (the variance estimated from the sample), one obtains [1, s.102]:

$$E(\hat{A}_j^2) \cong \frac{4}{n(j-1)} \sum_{i=1}^n (y(t_i) - \bar{y})^2 \quad [9]$$

where \bar{y} is the estimate of \hat{a}_0 , i.e. the arithmetic mean calculated from the sample.

The probability that an event that \hat{A}_j^2 will be α times greater than A_E^2 is determined by the pattern [1, s.102]:

$$P(\hat{A}_j^2 > \alpha A_E^2) = e^{-\alpha} \quad [10]$$

which – after rearrangement – gives [1, s.103]:

$$P\left\{\left(\frac{\hat{A}_j}{A_E}\right)^2 > -\ln \alpha\right\} = \alpha \quad [11]$$

If the level of significance presumed equals α one can verify a null hypothesis stating that the i-th jump that is the value of the quotient $\left(\frac{\hat{A}_j}{A_E}\right)^2$ is significant. If the inequality holds [1, s.103]

$$\left(\frac{\hat{A}_j}{A_E}\right)^2 > -\ln \alpha \quad [12]$$

then the verified null hypothesis should be rejected. This means that we can presume that the time series shows important oscillations with the period that equals n/i . This regularity holds with the probability $1 - \alpha$.

In some cases, information that some periodic oscillations are significant is enough to take the proper decision in relation to the source of the observed data. But in some other cases it is not enough; for instance, we need to predict what the probable course of the process observed in the near future will be. Let us ignore for the time being what the prognosis means and let us conduct our consideration on a decomposition of a time series a little further. Here a more advanced approach will be presented that makes use of the fundamental monograph written by Box and Jenkins (1976).

If all of the significant fluctuations are identified, the cyclic component function $C(t)^{16}$ can be constructed. Then, having specified two deterministic components, the third element can be identified, a purely stochastic one. The following sequence should be consider:

$$u(t_i) = x(t_i) - S(t_i) - C(t_i) ; i = 1, 2, \dots, n \quad [13]$$

The above relationship determines the time series of the residuals generated by the differences. It can be presumed that this sequence is a realisation of the unknown purely stochastic component \square_t . Usually, it is also presumed that the sequence $u(t_i)$ has a zero expected value and the finite and constant standard deviation \square_\square . In the majority of cases, a stronger presumption is formulated namely: $N_{\square}(0, \square_\square)$. Ignoring how strong the presumptions are, it is necessary to verify – by applying the appropriate statistical tests – whether all of these assumptions hold when confronted with the properties of data in hand.

In Figure 1 is a graph of a mass of hard coal sold in Poland versus time counted in months. The data comprises the period of the turn of the century: the late nineties – beginning of 21st century. A great restructuring of the mining industry (mainly coal) took place in Poland during this period.

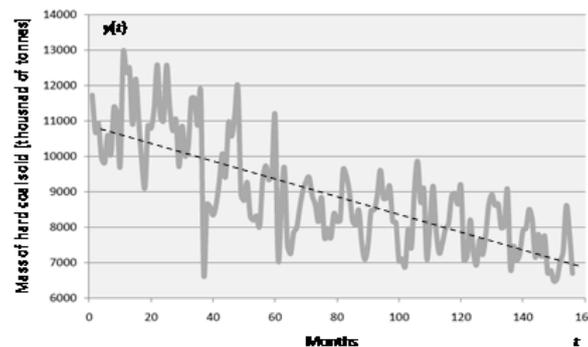


Fig. 1 Mass of hard coal sold versus time.

It is easy to notice that the time series is a non-stationary one and that it decreases with time. Therefore, a linear function¹⁷ was applied as the first approximation of the trend in the data, i.e.

$$\hat{y}_S(t) = \alpha t + \beta + u_t \quad [14]$$

The classical method of least squares was applied in order to estimate the unknown parameters α and β obtaining the following equation:

$$\hat{y}_S(t) = -25.56t + 10947 + \zeta_t \quad [15]$$

where ζ_t is a random component. The linear regression function is visible in this figure no 1.

It was suspected that there was a cyclic component in the sequence, thus ζ_t consisted of the constituent and the purely random one.

As the next step, the differences were calculated: $y(t) - \hat{y}_S(t)$ for all of the months of the sample. All of these differences are shown in figure no 2.

The sequence visible in this figure was the income data for the Fourier analysis with its transform applied. The Matlab 5.3 program was used and the result of the Fourier study is presented in figure no 3.

¹⁶ A cyclic component can consist of a few functions

¹⁷ This function makes sense in the interval observed and perhaps, for only a few months ahead. The more proper one should be a decreasing function that tends to a certain horizontal asymptote.

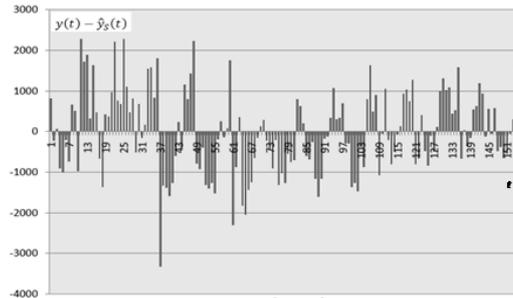


Fig. 2 Time series of the differences $y(t) - \hat{y}_S(t)$ (Manowska 2010).

It is important whether a dominating frequency is observed in terms of the module of the spectrum in such a graph. If so, a cyclic component is likely to exist.

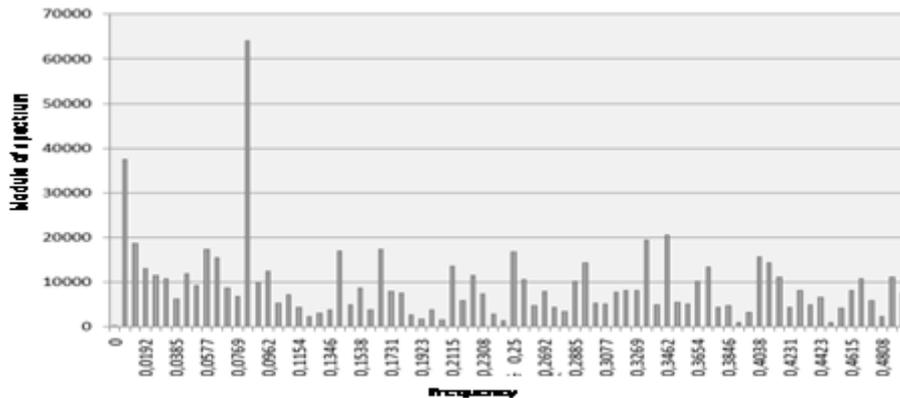


Fig. 3 Result of the application of the Fourier transform (Manowska 2010).

It was presumed that the general form of a cyclic function is described by the following pattern:

$$y_C(t) = \sin\left(2\frac{\pi}{T_0}t + \varphi\right)A \tag{16}$$

where: T_0 – period, number of months in one cycle,
 φ – the phase displacement,
 A – the amplitude.

Next, the harmonic analysis was applied that relies on a description of the residuals as the sum of the sinus functions for whatever period, amplitude and phase were selected using the appropriate algorithm. This algorithm relies on the analysis of the mean square error that is the result of the application of the sinus functions. The minimum of this error was a point of interest. In order to search for this minimum, the period was changed from 1 to 156 (sample size), the amplitude varied from 0 to 1000 tonnes and the phase displacement varied from 0 to 2π .

The minimum of the mean square error was 10,400 tonnes for the period equal to 12 and the phase displacement was 2.17. The amplitude was 820 tonnes. Therefore, the formula for the cyclic function was:

$$y_{C1}(t) = \sin\left(2\frac{\pi}{12}t + 2.17\right)820 \tag{17}$$

If so, a further point of interest was the next time series, the second residuals. They were obviously determined by the general formula: $y(t) - \hat{y}_S(t) - y_{C1}(t)$. This sequence is shown in figure 4.

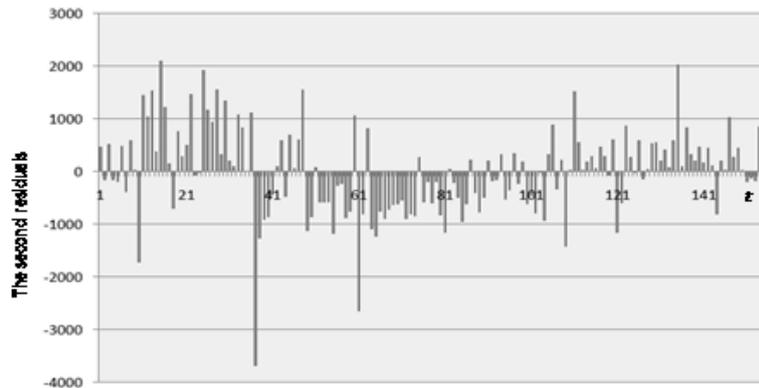


Fig. 4 Time series of the differences: $y(t) - \hat{y}_S(t) - y_{C1}(t)$ (Manowska 2010).

This way of further reasoning was repeated and the Fourier analysis applied. The result of its application is presented in Figure 5.

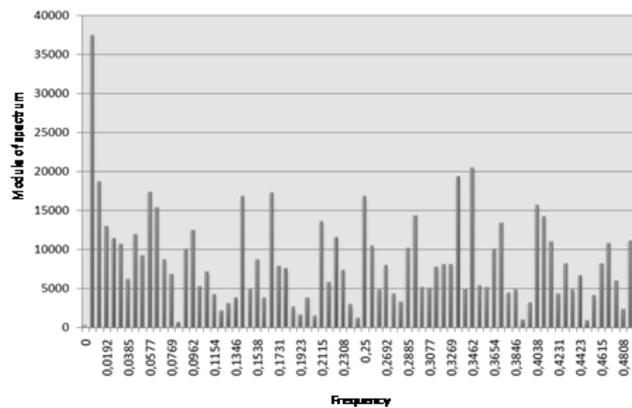


Fig. 5 Result of the application of the Fourier transform for the second residuals of the time series (Manowska 2010).

The minimum of the mean square error was 8,272 tonnes for the period equal to 114 and the phase displacement 0.72. The amplitude was 506 tonnes. Therefore, the formula for the cyclic function was:

$$y_{C2}(t) = \sin\left(2\frac{\pi}{114}t + 0.72\right) 506 \quad [18]$$

A further point of interest was the next time series, the third residuals, given as the difference: $y(t) - \hat{y}_S(t) - y_{C1}(t) - y_{C2}(t)$, for all points of the noted. This sequence is shown in figure 6.

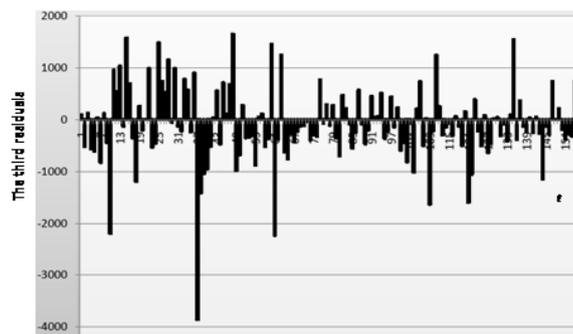


Fig. 6 Time series of the differences: $y(t) - \hat{y}_S(t) - y_{C1}(t) - y_{C2}(t)$ (Manowska 2010).

Based on this data, a third plot was constructed that showed the result of the application of the Fourier analysis. Figure 7 illustrates again the relationship between the modules of spectrum against time.

By analysing Figure 7 more carefully it is easy to conclude that there is no dominating frequency that is clearly greater than the others. Therefore, it can be suspected that no other cyclic functions should be constructed. However, this way of reasoning was repeated as in the previous cases and the mean square error was greater than before. It is obvious that the construction of the next cyclic function is not needed; a model with such a function will be worse than without it.

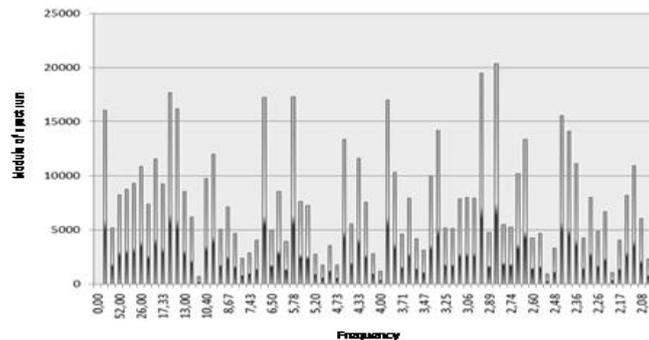


Fig. 7 Result of the application of the Fourier transform for the third residuals of the time series (Manowska 2010).
Thus, the final model to generate the observed time series is:

Thus, the final model to generate the observed time series is:

$$y(t) = \hat{y}_S(t) + y_{C1}(t) + y_{C2}(t) + \xi_t \quad [19]$$

III. Conclusion

The demand for coal is the size of a very difficult planning and forecasting, especially at a time of economic instability and the production of basic industries (Karbownik, Turek, 2000). The topic of market analysis Polish coal can be found in the publications of many authors (see, eg, Karbownik Wodarski, 2005; Karbownik Bijańska, 2000, Bendkowski, Przybyła, 1999). Moreover, you need to consider the fact the changes currently taking place in the global market. These changes are not for the Polish preferred. The price of coal will be declined in the longer term. This is influenced by a decline in U.S. coal prices (due to increased production from shale). Influence will also increase in coal production in Indonesia and Australia, the planned increase in production in Colombia (32 million tonnes per annum to 60 million tonnes), and the fact of finding rich, shallowly buried deposits in Mongolia. Sales forecasting coal is essential for any manufacturer, supplier and retailer. Forecasts of demand in the future will determine the amount to be manufactured and delivered. The mining company can not wait for the sale, and then respond to it. Companies must anticipate and plan for future needs so that they can respond immediately to customer orders. Thus, companies are forced to operate well in advance to plan their investments. Other hand easy access to other energy sources such as oil, natural gas, etc. are the main causes of "overproduction" of coal. This forces the use of extractive industries advanced techniques identify potential sale of coal, analyzes the elasticity of demand with respect to price, as well as evaluating the behavior of competitors.

References

- [1]. Abbaspour M., Karbassi A., Khalaji Asadi M, Moharamnejad N, Khadivi S., Ali Moradi M.: Energy Demand Model of the Household Sector and Its Application in Developing Metropolitan Cities (Case Study: Tehran), Journal of Environmental of studies, Vol. 22, No. 2 (2013), pages 319-329, 2013
- [2]. Box G. E. P., Jenkins G.M.: Analiza szeregów czasowych. Prognozowanie i sterowanie. Państwowe Wydawnictwo Naukowe, Warszawa 1983
- [3]. Boyce G.: Meeting the challenges of a New energy era". CoalTrans Berlin 2002
- [4]. BP Statistical Review of World Energy, June 2007
- [5]. Carbon reduction opens up producers opportunities W: Coal Trans International, April 2004,
- [6]. Czaplicki J.M.: Statistics for Mining Engineering, CRC Press, London 2014.
- [7]. China Coal Report 2005, 2006
- [8]. Coal information IEA/OECD Paris 2000 – 2006
- [9]. Euracoal Market Raport 1/2007"
- [10]. Fukata K., Washio T., Yada K., Motoda H.: A Method to Search ARX Model Orders and Its Application to Sales Dynamics Analysis, The Institute of Scientific and Industrial Research, Osaka University 2006
- [11]. Generation Capacity Planning, TSO: A New methodology for forecasting long term electricity dem and for the republic Ireland, 2002
- [12]. Juan R Trapero, Nikolaos Kourentzes & Robert Fildes: On the identification of sales forecasting models in the presence of promotions, Journal of the Operational Research Society (2014)doi:10.1057/jors.2013.174
- [13]. Manowska A.: Prognozowanie wielkości sprzedaży węgla kamiennego dla grupy kopalń, rozprawa doktorska, Politechnika Śląska, pages 170-175, 2010
- [14]. Manowska A.: Planowanie wielkości produkcji na podstawie statystycznej analizy szeregu czasowego wielkości sprzedaży węgla kamiennego na rynku krajowym, Wydaw. Akademii Techniczno-Humanistycznej, pages 267-282, 2012
- [15]. Mohr S.H., Evans G.M.: Forecasting coal production until 2100, Fuel, Volume 88, Issue 11, November 2009, pages 2059–2067
- [16]. Paszcza H., Pitura K.: Górnictwo węgla kamiennego w Polsce w 2012 roku, Wiadomości Górnicze 5/2013.
- [17]. Patzek T., Croft G: A global coal production forecast with multi-Hubbert cycle analysis, EnergyVolume 35, Issue 8, August 2010, pages 3109–3122
- [18]. Ritschel W. SchiffeH. W.: World Market for Hard Coal, 2013
- [19]. Raport zbiorczy: Zintegrowany program rozwoju źródeł energii elektrycznej do 2020 roku (ZPR – 2) Warszawa 1997r.,
- [20]. Sablik J.: Model ekologicznego i ekonomicznego prognozowania wydobycia i użytkowania czystego węgla. Tom 1, Główny Instytut Górnictwa, Katowice 2004,

- [21]. Schiffer H.W: Długofalowe bezpieczeństwo w dostawach energii – udział węgla” Mat. Kof. „Przyszłość węgla w gospodarce światowej i Polski” Katowice 2004,
- [22]. Secomski K.: Podstawowe elementy studiów prognostycznych. Wykład o podstawach prognozowania, PAN, Komitet Badań i Prognoz „Polska 2000”, Warszawa 1970,
- [23]. Strategia działalności górnictwa węgla kamiennego w Polsce w latach 2007 – 2015, Warszawa 2007