Appeared in: Physica A 336 (2004) pp. 196-205

On detecting and modeling periodic correlation in financial data

E. Broszkiewicz-Suwaj [a,1], A. Makagon [b], R. Weron [c,1,2], A. Wyłomańska [a,1]

[a] *Institute of Mathematics, Wrocław University of Technology, 50-370 Wrocław, Poland*

[b] *Department of Mathematics, Hampton University, Hampton, VA 23668, USA*

[c] *Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology, 50-370 Wrocław, Poland*

Abstract

For many economic problems standard statistical analysis, based on the notion of stationarity, is not adequate. These include modeling seasonal decisions of consumers, forecasting business cycles and – as we show in the present article – modeling wholesale power market prices. We apply standard methods and a novel spectral domain technique to conclude that electricity price returns exhibit periodic correlation with daily and weekly periods. As such they should be modeled with periodically correlated processes. We propose to apply periodic autoregression (PAR) models which are closely related to the standard instruments in econometric analysis – vector autoregression (VAR) models.

Key words: periodic correlation, sample coherence, electricity price, periodic autoregression, vector autoregression
PACS: 02.70.Hm, 05.45.Tp, 89.65.Gh

1 Introduction

Conventional time series analysis is heavily dependent on the assumption of stationarity. But this assumption is unsatisfactory for many physical processes of interest. Periodically correlated (PC) processes offer an alternative. They

[1] Research partially supported by KBN Grant PBZ-KBN 016/P03/99.
[2] Corresponding author. E-mail address: rweron@im.pwr.wroc.pl.

are non stationary but possess many of the properties of stationary processes. Hence, the numerous attempts to apply PC processes in various areas of science and technology.

The theory of periodically correlated processes, begun by Gladyshev [1], has developed along probabilistic lines and the emphasis has been on the underlying structure (for a review see [2,3]). However, significant consideration has been also given to applications. A particularly interesting fact about PC processes, which will be exploited in the last Section of this paper, is that they can be viewed as stationary processes that take values in a larger space. For example, finite variance PC sequences of period T are equivalent to T-dimensional stationary vector sequences.

Periodically correlated processes exhibit a periodic "rhythm" that is generally much more complicated than periodicity in the mean (which is a manifestation of the classical notion of periodicity). This is due to the fact that for a stochastic sequence to be periodically correlated its autocovariance function $R_X(m,n) = E[(X_m - \mu_X(m))\overline{(X_n - \mu_X(n))}]$ has to be periodic with the same period T as its mean, i.e. $R_X(m,n) = R_X(m+T, n+T)$ and $\mu_X(n) = E(X_n) = \mu_X(n+T)$ for all integers m and n. This "rhythm" has led to the application of PC processes in such diverse disciplines as climatology [4–6], hydrology [7–9], electrical engineering [10], signal processing [11,12] and economics [13–18]. In this paper we pursue the latter application further and study the periodic structure of electricity prices. Before we go on with the empirical analysis, in the next Section we describe the methods for detecting periodicity.

2 Methods for detecting periodicity

In the frequency domain the standard tool for detecting periodicity is the sample analogue of the spectral density – the periodogram. It is defined as

$$P_N(\omega_j) = \frac{1}{2\pi N}|I_N(\omega_j)|^2,$$

where $I_N(\omega_j)$ is the discrete Fourier transform $I_N(\omega_j) = \sum_{n=0}^{N-1} X_n e^{-2\pi i \omega_j n}$, $\{X_1, \ldots, X_N\}$ is the analyzed sample and $\omega_j = j/N$, $j = 1, \ldots N$ are the frequencies. For many PC processes the periodogram correctly detects periodicities, but there are also many for which it fails. For example, consider a PC process of the form $X_n = S_n \cdot f^2(n)$, where $f(n) = \mathrm{mod}(n, 8)$, i.e. division modulo 8, $S_n = -0.3S_{n-1} + 0.4S_{n-2} + \epsilon_n$ is a stationary time series and $\{\epsilon_n\}$ denotes a white noise sequence. The periodogram does not exhibit any regularity and only methods based on sample coherence (defined below) detect the proper period of $T = 8$, see Fig. 1.

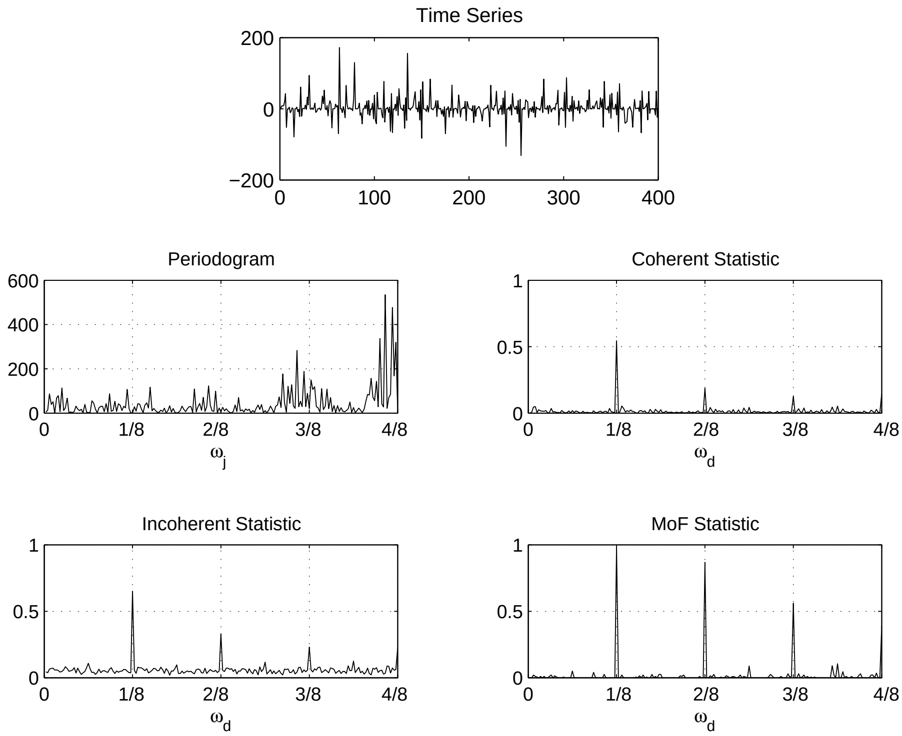

Fig. 1. A sample of $N = 400$ points (*top panel*) and tests for periodic correlation of a PC process of the form $X_n = S_n \cdot f^2(n)$, where $f(n) = \mathrm{mod}(n, 8)$, $S_n = -0.3S_{n-1} + 0.4S_{n-2} + \epsilon_n$ and $\{\epsilon_n\}$ denotes a white noise sequence. The following test parameters were used: M=20, B=100 and α=0.01.

Since the classical frequency domain method often lets us down we have to seek better techniques for detecting periodic correlation. A good starting point is the comparison of the 2-dimensional spectral density structures of stationary and of periodically correlated processes, see Fig. 2. It can be shown [19–21] that the spectral density of the latter is characteristically constrained to be in the intersection of the $2T-1$ diagonal lines $\{(2\pi\omega_p, 2\pi\omega_q) : \omega_q = \omega_p - \frac{k}{T}, k = -(T-1), \ldots, T-1\}$ with the square $[0, 2\pi) \times [0, 2\pi)$. In case of stationary sequences the support reduces to the diagonal of this square $\{\omega_q = \omega_p\}$. For this reason we can reduce the 2-dimensional density of a stationary sequence to the 1-dimensional one and estimate it using the periodogram. The same property is responsible for the failure of the periodogram for many PC processes and at the same time can provide us with techniques for testing periodic correlation.

To investigate the presence of periodic correlation we can use a very serviceable statistic which is a smoothed estimator of the normalized, 2-dimensional spectral density function. It is called sample coherence and is defined as

$$|\hat{\gamma}(p, q, M)|^2 = \frac{|\sum_{m=0}^{M-1} I_N(\omega_{p+m})\overline{I_N(\omega_{q+m})}|^2}{\sum_{m=0}^{M-1} |I_N(\omega_{p+m})|^2 \sum_{m=0}^{M-1} |I_N(\omega_{q+m})|^2}, \quad (1)$$

where $0 < p, q \leq N$, N is the sample length and M is the smoothness co-

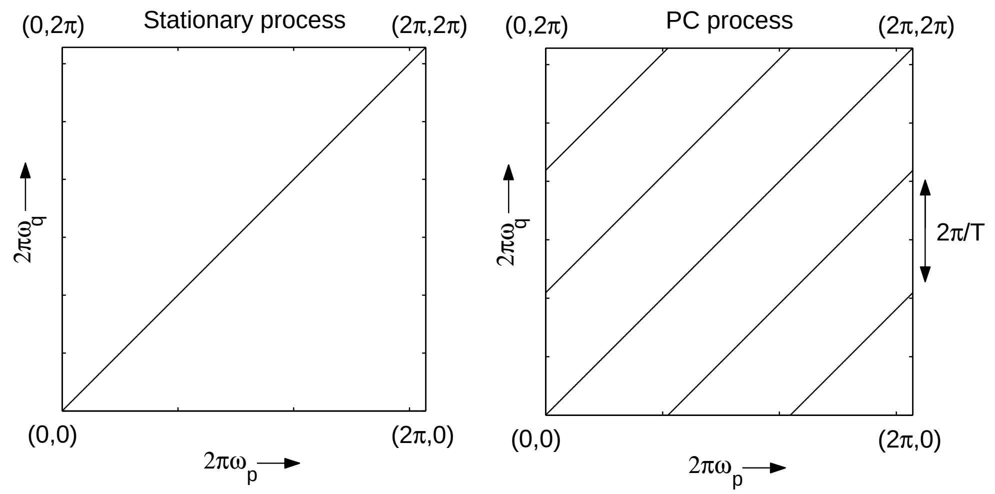

Fig. 2. Spectral mass location for a stationary and a periodically correlated (PC) process.

efficient. Sample coherence takes only real values between 0 and 1. For PC processes the values taken on the support lines are significantly different from those taken for the intermediate frequencies. This property yields "graphical" methods for detecting periodic correlation to be discussed in the next paragraph.

2.1 The coherent and incoherent statistics

Hurd and Gerr [22] proposed two tools for determining the presence of periodic correlation – the coherent and the incoherent statistics. The former is defined as

$$|\hat{\gamma}(0, d, N)|^2,$$

i.e. the sample coherence, given by formula (1), evaluated for $M = N$, whereas the latter is given by

$$\delta(d, M) = \frac{1}{L+1} \sum_{p=0}^{L} |\hat{\gamma}(pM, pM + d, M)|^2,$$

where $L = [\frac{N-1-d}{M}]$ and $d = |q - p|$. Since both statistics depend on the differences between frequencies, the plots against d (or ω_d) are the most indicative. Moreover, the statistics are usually plotted only in the interval $(0, \frac{N}{2})$, because the values in the interval $(\frac{N}{2}, N)$ are a mirror image of the values in the former one. Peaks at points $\omega_d, \omega_{2d}, \omega_{3d}$, etc. indicate periodic correlation with period of length $T = \frac{1}{\omega_d}$. Note that both methods do not detect the exact period of the process but the periods of its harmonic components. As a result we may see more than one spike, see Fig. 1. The coherent and incoherent statistics are,

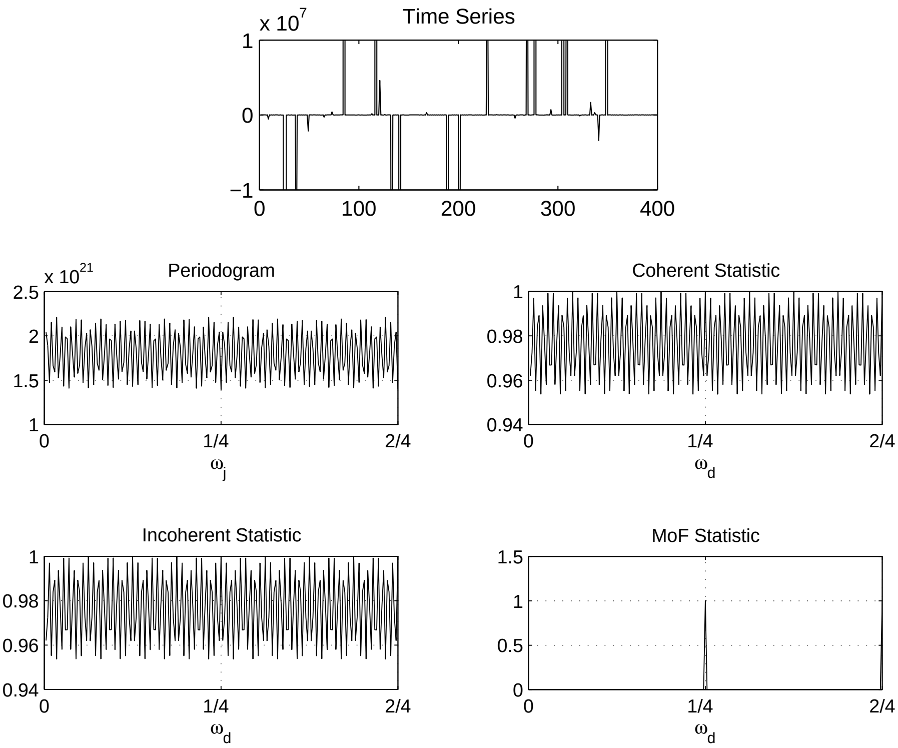

Fig. 3. A sample of $N = 400$ points (*top panel*) and tests for periodic correlation of a PC process of the form $X_n = S_n \cdot \exp[8 \cdot \{1 + \sin(\frac{\pi n}{2}) \cdot \frac{4}{5}\xi_n\}]$, where $S_n = -0.3S_{n-1} + 0.4S_{n-2} + \frac{1}{2}\eta_n$ is a stationary time series; $\{\xi_n\}$ and $\{\eta_n\}$ denote independent white noise sequences. The test parameters used are the same as in Fig. 1, i.e. M=20, B=100 and α=0.01. Only the MoF statistic returns the true period of $T = 4$.

in general, much better than the periodogram. However, it is still pretty easy to find examples of PC processes for which these techniques fail, see Fig. 3.

2.2 The MoF statistic

To cope with the moderate efficiency of the coherent and incoherent statistics a new technique was proposed recently [23]. The measure of fitness (MoF) statistic is based on the bootstrap methodology and is defined by

$$\mathrm{MoF}(d, M) = \frac{1}{N}\sum_{p=1}^{N} \kappa_\alpha(p, p+d, M),$$

where

$$\kappa_\alpha(p, q, M) = \begin{cases} 1 & |\hat{\gamma}(p, q, M)|^2 \geq \hat{c}_\alpha, \\ 0 & |\hat{\gamma}(p, q, M)|^2 < \hat{c}_\alpha, \end{cases}$$

α is the confidence level and $\hat{c}_\alpha$ is the estimator of the critical value, which is computed using the Moving Blocks Bootstrap (MBB) procedure [24,25]. Like the coherent and incoherent statistics, the MoF statistic takes real values in the interval $[0, 1]$ and – due to the symmetry – is plotted only in the interval $(0, \frac{N}{2})$. Peaks at points $\omega_d, \omega_{2d}, \omega_{3d}$, etc. indicate periodic correlation with period of length $T = \frac{1}{\omega_d}$. What distinguishes the MoF statistic from the former two is the summation scheme in which only significant (at some level) values of sample coherence are used. It is not the value of sample coherence that is important but its value relative to values at other frequencies. Thus the MoF statistic detects periodic correlation even for processes exhibiting extreme volatility, see Fig. 3.

3 Electricity prices

Having described methods for detecting periodicity we are ready to apply them to market data. The analyzed data set contains spot prices from the Nord Pool [26] power exchange for delivery of electricity in the Nordic region for every hour since January 1st, 1997 until December 31st, 2001. It was kindly provided by SKM Market Predictor AS.

Due to the nature of electricity trade, spot prices exhibit a behavior not observed in other financial or commodity markets. Climate and working hours driven seasonal fluctuations in demand translate into seasonal behavior of spot prices [27–29]. In addition, limited flexibility of the supply side coupled with outages, transmission limitations and extreme weather conditions cause the spot electricity prices to exhibit infrequent, but large jumps and an anti-persistent behavior with the Hurst exponent ranging from 0.25 to 0.42 [30–32].

In this paper we ask ourselves whether electricity spot prices are governed by a PC process. To this end we investigate the returns of spot prices. Due to the computational complexity of the MoF technique we split the analysis into two steps. First we test all five years of daily data (260 weeks, 1820 data points) with the returns calculated from average daily prices, see Fig. 4. As it turns out, all methods detect a 7 day period; peaks appear at frequencies being multiples of $\frac{1}{7}$. This is in agreement with earlier observations [27,33]. In the next step we zoom in to see the dependencies at a finer time scale and analyze 100 days (January 1st – April 10th, 1997) of hourly data, see Fig. 5. Again all methods detect a distinct period, this time of $T = 24$ hours (peaks appear at frequencies being multiples of $\frac{1}{24}$).

The combined two step analysis leaves us with a clear signal that electricity spot price returns are periodically correlated. Hence, we should either deseasonalize the data with respect to the daily and the weekly cycle and then

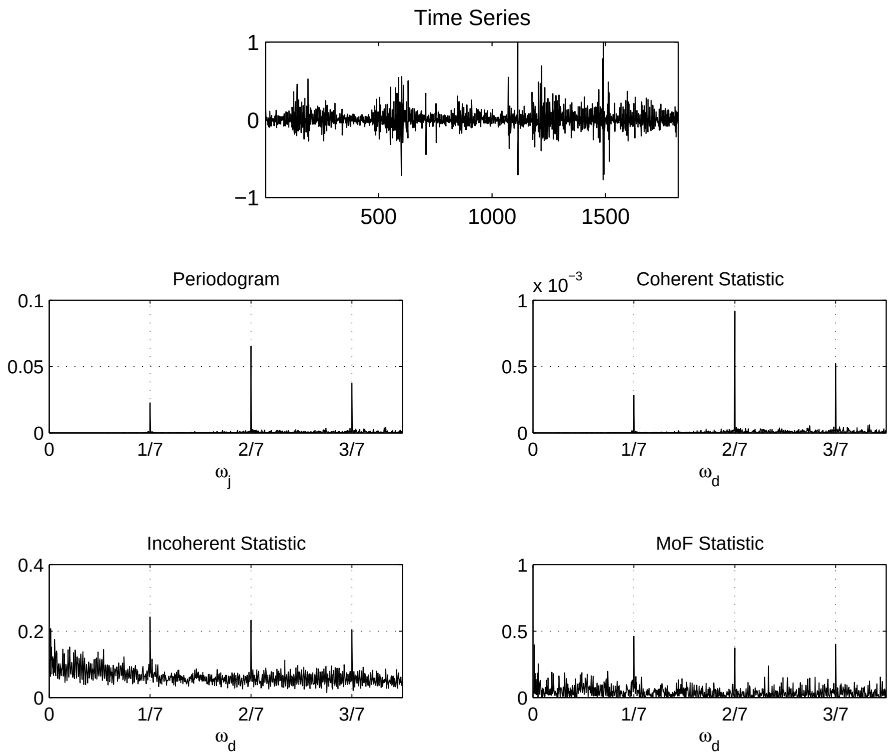

Fig. 4. Daily returns of Nord Pool electricity spot prices during the period January 1st, 1997 – December 25th, 2001 (*top panel*) and plots of all statistics for these returns (260 weeks, 1820 data points). Returns were calculated based on average daily prices. The test parameters used are the same as in Figs. 1 and 3, i.e. M=20, B=100 and α=0.01.

model the residuals via a stochastic process (like in [28,29]) or seek models of electricity price dynamics in the class of PC processes. In the next Section we follow the latter approach and describe a prominent class of PC models.

4 Periodic autoregression models

As McLeod [34] observes, periodically correlated series should not be modeled with the widely used in econometrics seasonal autoregressive moving-average (SARMA) class. This is because SARMA models, contrary to their name, are actually stationary models with large (in absolute value) autocovariances at lags that are multiples of the period. A flexible class of models that have the desirable properties is the class of periodic autoregressive moving-average (PARMA) models. Analogous to ARMA models and short memory stationary series, PARMA models are fundamental periodic and periodically correlated time series models [8,17,35].

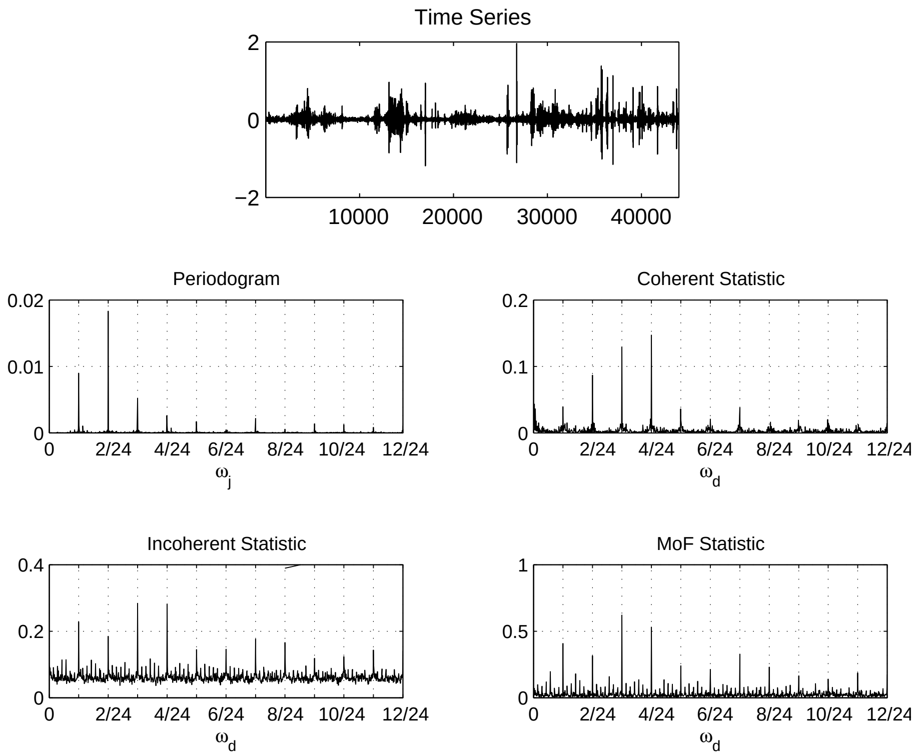

Fig. 5. Hourly returns of Nord Pool electricity spot prices during the period January 1st, 1997 – December 25th, 2001 (*top panel*) and plots of all statistics for hourly returns since January 1st until April 10th, 2001 (100 days, 2400 data points). The test parameters used are the same as in Figs. 1, 3 and 4, i.e. M=20, B=100 and α=0.01.

A PARMA(p, q) system is defined by

$$x(n) - \sum_{k=1}^{p} \phi_k(n) x(n-k) = \sum_{k=1}^{q} \theta_k(n) \xi_{n-k} + \xi_n, \tag{2}$$

where the scalar sequences $\{\phi_k(n)\}$ and $\{\theta_k(n)\}$ are periodic in n with the same period $T \geq 1$ and $\{\xi_n\}$ is a white noise sequence. Observe that $T = 1$ reduces the model to a classical ARMA system. Periodic moving-average processes do not appear to be useful in economics [17]. Hence, the econometric analysis of periodically correlated time series concentrates on periodic autoregressive (PAR) processes of the form

$$x(n) - \sum_{k=1}^{p} \phi_k(n) x(n-k) = \xi_n. \tag{3}$$

The key to the deeper analysis of periodic and periodically correlated processes lies in adoption of a vector representation [8,36]. PAR models are associated to vector autoregression (VAR) models – the two models are parallel in the sense that there is a one-to-one correspondence between solutions of the systems. To see this arrange coefficients of the sum in formula (3) into a $T \times (l+1)T$ matrix with $l \geq 1 + p/T$

$$\begin{bmatrix} 1 & -\phi_1(0) & -\phi_2(0) & \ldots & -\phi_p(0) & 0 & 0 & \ldots & 0 \\ 0 & 1 & -\phi_1(1) & \ldots & -\phi_{p-1}(1) & -\phi_p(1) & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & -\phi_{p-2}(2) & -\phi_{p-1}(2) & -\phi_p(2) & \ldots & 0 \\ & \vdots & & & & \vdots & & & \\ 0 & 0 & 0 & \ldots & 1 & -\phi_1(T-1) & \ldots & -\phi_p(T-1) & \ldots \end{bmatrix}$$

and denote the consecutive $T \times T$ blocks of the matrix by $\Phi_0, \ldots, \Phi_l$, then (3) can be written as the T-dimensional VAR system

$$X(n) - \sum_{k=1}^{l} \Phi_k X(n-k) = \Xi_n, \tag{4}$$

where the column vectors $X(n)$ and Ξ_n are given by $X(n) = [x(nT), x(nT-1), \ldots, x((n-1)T+1)]'$ and $\Xi_n = [\xi_{nT}, \xi_{nT-1}, \ldots, \xi_{(n-1)T+1}]'$. System (3) has a unique periodically correlated solution if and only if (4) has a unique stationary solution. The same result holds for PARMA and VARMA systems. Therefore the analysis of PAR (PARMA) models reduces to the examination of VAR (VARMA) systems. Luckily, a lot is known about the latter as they have been extensively studied in econometrics since the pioneering work of Sims [37] in 1980. Vector autoregression models are by now standard instruments in econometric analysis and parameter estimation schemes are well known [38–40]. However, a lot of technical issues arise when analyzing data sets of a few thousand observations.

5 Conclusions

Periodic and periodically correlated processes have not been very widely applied in economics to date. Nevertheless, a number of studies show that periodic and PC processes can arise naturally from the application of economic theory to modeling decisions in economic context and their role should not be dismissed as unimportant. Examples include Osborn [14] who argues that a process of periodic structure arises when modeling seasonal decisions of consumers, Hansen and Sargent [15] who suggest that PC processes could also arise from seasonal technology, and Ghysels [41] who explores the periodic nature of U.S. business cycle turning points. Despite their applicability, prediction and likelihood evaluation methods for generic PC processes remain relatively unexplored [42]. Fortunately, estimation techniques exist for the special case of PAR (PARMA) models discussed in Section 4.

In the present paper we have looked at the power market and studied the behavior of spot electricity prices as traded at the Nord Pool power exchange. By applying spectral domain techniques (including the novel MoF statistic [23]) we concluded that electricity price returns exhibit periodic correlation with daily (24 hours) and weekly (7 days or 168 hours) periods. As such they could be modeled with PC processes like the periodic autoregression (PAR) system. Although the general estimation scheme, through vector autoregression (VAR) models, is known a lot of technical issues remain unsolved and will be the subject of further research.

References

[1] E.G. Gladyshev, Soviet Math. 2 (1961) 385.

[2] W.A. Gardner, Statistical Spectral Analysis: A Nonprobabilistic Theory, Prentice Hall, Englewood Cliffs, NJ, 1987.

[3] A.M. Yaglom, Correlation Theory of Stationary and Related Random Functions, Springer-Verlag, New York, 1987.

[4] E.J. Hannan, Australian J. Physics 8 (1955) 289.

[5] P. Bloomfield, H.L. Hurd, R. Lund, J. Time Series Anal. 15 (1994) 127.

[6] R. Lund, H.L. Hurd, P. Bloomfield, R.L. Smith, J. Climate 11 (1995) 2787.

[7] H.A. Thomas, M.B. Fiering, Mathematical synthesis of stream flow sequences for the analysis of river basins by simulation, in *Design of Water Resources*, ed. A. Maass et al., Harvard University Press, 1962.

[8] A.V. Vecchia, Water Resources Bull. 21 (5) (1985) 721.

[9] A.I. McLeod, J. Time Series Anal. 15 (1994) 221.

[10] W.A. Gardner, L.E. Franks, IEEE Trans. Information Theory 21 (1975) 4.

[11] W.R. Bennett, Bell System Tech. J. 37 (1958) 1501.

[12] W.A. Gardner (ed.), Cyclostationarity in Communications and Signal Processing, IEEE Press, 1994.

[13] M. Pagano, E. Parzen, J. Econometrics 9 (1979) 137.

[14] D.R. Osborn, J. Applied Economics 3 (1988) 255.

[15] L.P. Hansen, T.J. Sargent, J. Econometrics 55 (1993) 21.

[16] P.H. Franses, Periodicity and Stochastic Trends in Economic Time Series, Oxford University Press, 1996.

[17] E. Ghysels, D.R. Osborn, The Econometric Analysis of Seasonal Time Series, Cambridge University Press, 2001.

[18] P.H. Franses, R. Paap, Periodic Time Series Models, Oxford University Press, 2004.

[19] A. Makagon, A.G. Miamee, H. Salehi, Periodically correlated processes and their spectrum, in *Nonstationary Stochastic Processes and Their Applications*, ed. A.G. Miamee, World Scientific, 1992, pp. 147-164.

[20] D. Dehay, H.L. Hurd, Representation and estimtaion for periodically and almost periodically correlated random processes, in [12], 1994.

[21] A. Makagon, J. Multivariate Anal. 78.1 (2001) 1.

[22] H.L. Hurd, N.L. Gerr, J. Time Series Anal. 12 (1991) 337.

[23] E. Broszkiewicz-Suwaj, Methods for determining the presence of periodic correlation based on the bootstrap methodology, Hugo Steinhaus Center Research Report HSC/03/2, 2003. See:
http://www.im.pwr.wroc.pl/˜hugo/Publications.html.

[24] D. Politis, J. Romano, M. Wolf, Subsampling, Springer, New York, 1999.

[25] J. Leśkow, K. Żychaluk, Resampling in assessing stationarity of econometric time series, WSB-NLU preprint.

[26] Nord Pool Annual Report, Nord Pool ASA, 2002.

[27] V. Kaminski (ed.), Managing Energy Price Risk, Risk Books, London, 1999.

[28] J.J. Lucia, E.S. Schwartz, Rev. Derivatives Research 5 (2002) 5.

[29] R. Weron, I. Simonsen, P. Wilman, Modeling highly volatile and seasonal markets: evidence from the Nord Pool electricity market, in *The Application of Econophysics*, ed. H. Takayasu, Springer, Tokyo, 2004, pp. 182-191.

[30] R. Weron, B. Przybyłowicz, Physica A 283 (2000) 462.

[31] R. Weron, Measuring long-range dependence in electricity prices, in *Empirical Science of Financial Fluctuations*, ed. H. Takayasu, Springer, Tokyo, 2002, pp. 110-119.

[32] I. Simonsen, Physica A 322 (2003) 597.

[33] R. Weron, Physica A 285 (2000) 127.

[34] A.I. McLeod, Int. Stat. Rev. 61 (1993) 187.

[35] R.B. Lund, I.V. Basawa, J. Time Ser. Anal. 20 (2000) 53.

[36] M. Pagano, Ann. Statist. 6 (1978) 1310.

[37] C.A. Sims, Econometrica 48 (1980) 1.

[38] H. Lütkepohl, Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin, 1991.

[39] J.D. Hamilton, Time Series Analysis, Princeton University Press, Princeton, 1994.

[40] A. Benkwitz, Multiple Time Series Analysis, in *XploRe Application Guide*, eds. W. Härdle, Z. Hlávka, S. Klinke, Springer, Berlin, 2000, pp. 459-482.

[41] E. Ghysels, J. Business Econ. Statist. 12 (1994) 289.

[42] A. Makagon, Probab. Math. Statist. 19.2 (1999) 287.