DYNAMIC ECONOMETRIC MODELS
Vol. 6 – Nicholas Copernicus University – Toruń – 2004

Liliana Talaga
University of Szczecin

Predictors of Non-Stationary ARIMA Processes

This article contains a comparison of stochastic properties of non-stationary ARIMA(*p*,*k*,*q*) regular processes and their predictors by means of frequency characteristics of gain function and function of phase angle.

1. The general non-stationary regular process

The general model, which generates *regular stationary process* X_t is described by the following difference equation:

$$\sum_{s=0}^{p} \delta_s \cdot X_{t-s} = \sum_{r=0}^{q} \gamma_r \cdot \varepsilon_{t-r} , \tag{1.1}$$

where:
δ_s, γ_r – real constants (weights), ε_t – white noise.
By means of operators the expression (1.1) can be written as:

$$B(U) \cdot X_t = C(U) \cdot \varepsilon_t , \tag{1.2}$$

where:
$B(U)$ – stationary operator of autoregression of the *p*-th order,

$$B(U) = 1 - \beta_1 \cdot U - \beta_2 \cdot U^2 - ... - \beta_p \cdot U^p , \tag{1.3}$$

$-\beta_s = \delta_s$,
U – the backward operator,

$$U^s \cdot X_t = X_{t-s} , \tag{1.4}$$

$C(U)$ – the operator of moving average of the *q*-th order,

$$C(U) = 1 + \gamma_1 \cdot U + \gamma_2 \cdot U^2 + ... + \gamma_q \cdot U^q . \tag{1.5}$$

Process (1.2) is called an ARMA(p,q) process. It is stationary, if all roots of $B(U)=0$ equation lie outside the unit circle and it is invertible if the roots of $C(U)=0$ equation belong to the $|U|>1$.

The evolutional non-stationary processes can be written by means of the generalized operator of autoregression $B^*(U)$ for which *k-number of roots* of the $B^*(U)=0$ equation are equal to one and the others lie outside the unit circle.

$$B^*(U)=\Delta^k \cdot B(U), \tag{1.6}$$

where:

$B^*(U)$ – the non-stationary operator of (p+k)-th ordered autoregression,

$$B^*(U)=1-\beta_1^* \cdot U-\beta_2^*-...-\beta_{p+k}^* \cdot U^{p+k}, \tag{1.7}$$

Δ – the difference operator[1],

$$\Delta X_t = X_t - X_{t-1} = (1-U) \cdot X_t, \tag{1.8}$$

$$\Delta^n \cdot X_t = \Delta^{n-1} \cdot X_t - \Delta^{n-1} \cdot X_{t-1} = (1-U)^n \cdot X_t. \tag{1.9}$$

The general non-stationary regular process, also called ARIMA autoregressive-integrated moving average process of the (p,k,q)-th order can be written as:

$$X_t - \beta_1^* \cdot X_{t-1} - ... - \beta_{p+k}^* \cdot X_{t-p-k} = \varepsilon_t + \gamma_1 \cdot \varepsilon_{t-1} + ... + \gamma_q \cdot \varepsilon_{t-q}, \tag{1.10}$$

$$B^*(U) \cdot X_t = B(U) \cdot \Delta^k \cdot X_t = C(U) \cdot \varepsilon_t, \tag{1.11}$$

$$\Pi^*(U) \cdot X_t = \varepsilon_t, \tag{1.12}$$

where:

$\Pi^*(U)$ – the non-stationary operator of autoregression of order ∞,

$$\Pi^*(U) = 1 - \pi_1^* \cdot U - \pi_2^* \cdot U^2 - ... \tag{1.13}$$

$$\Pi^*(U) = \frac{B^*(U)}{C(U)}. \tag{1.14}$$

If all the roots of $C(U)=0$ polynomial are beyond the unit circle, the $\Pi^*(U)$ operator satisfies the following condition:

$$C(U) \cdot \Pi^*(U) = B^*(U). \tag{1.15}$$

ARIMA (p,k,q) model generates processes, where k-th difference is a stationary regular process.

[1] The difference operator Δ and backward U are linear, they are characterized by following equation: $\Delta = 1 - U$.

2. The predictor of ARIMA(p,k,q) process

With reference to (1.12) formula the ARIMA(p,k,q) process can be written as:

$$X_t = \pi_1^* \cdot X_{t-1} + \pi_2^* \cdot X_{t-2} + \ldots + \varepsilon_t . \tag{2.1}$$

Substituting $s = t + h$, we obtain an equation, which is the basis in the certain predictor of the process for h-step forecast, $h = 1,2,\ldots,H$:

$$X_{t+h} = \pi_1^* \cdot X_{t+h-1} + \pi_2^* \cdot X_{t+h-2} + \ldots + \varepsilon_{t+h} . \tag{2.2}$$

The best mean-squared predictor of the regular *component* for h-step forecast is described as follows:

$$\hat{X}_{t+h} = \pi_1^* \cdot \hat{X}_{t+h-1} + \pi_2^* \cdot \hat{X}_{t+h-2} + \ldots + \hat{\varepsilon}_{t+h} , \tag{2.3}$$

where:

$$\hat{X}_s = X_s, \quad \text{gdy} \quad s \le t,$$

$$\hat{\varepsilon}_{t+h} = 0.$$

The (2.3) equation can be also written as:

$$\hat{X}_{t+h} = \beta_1^* \cdot \hat{X}_{t+h-1} + \beta_2^* \cdot \hat{X}_{t+h-2} + \ldots + \beta_{p+k}^* \cdot \hat{X}_{t+h-(p+k)} + \hat{\varepsilon}_{t+h} + \\ + \gamma_1 \cdot \hat{\varepsilon}_{t+h-1} + \ldots + \gamma_q \cdot \hat{\varepsilon}_{t+h-q} , \tag{2.4}$$

where:

$$\hat{X}_s = X_s, \qquad \text{for} \quad s \le t$$

$$\hat{\varepsilon}_s = \begin{cases} 0, & \text{for} \quad s > t \\ \varepsilon_s, & \text{for} \quad s \le t. \end{cases}$$

The forecast error $\delta_t(h)$ is explained by the following equation:

$$\delta_t(h) = x_{t+h} - \hat{x}_t(h), \tag{2.5}$$

where:

x_{t+h} – the real value of time series in the forecasted period;

$\hat{x}_t(h)$ – the prognosis made in t-th moment with the h-step forecast;

h – the forecast of the time series (the forecast horizon), $h = 1, 2, \ldots, H$.

The forecast error (2.5) can be also written as:

$$\delta_t(h) = \pi_1^* \cdot \delta_t(h-1) + \pi_2^* \cdot \delta_t(h-2) + \ldots + \pi_{h-1}^* \cdot \delta_t(1) + \varepsilon_{t+h} . \tag{2.6}$$

From the expression (2.6) we obtain:

$$\delta_t(1) = \varepsilon_{t+1}$$
$$\delta_t(2) = \pi_1^* \cdot \delta_t(1) + \varepsilon_{t+2} = \varepsilon_{t+2} + \pi_1^* \cdot \varepsilon_{t+1}$$
$$\delta_t(3) = \pi_1^* \cdot \delta_t(2) + \pi_2^* \cdot \delta_t(1) + \varepsilon_{t+3} = \varepsilon_{t+3} + \pi_1^* \cdot \varepsilon_{t+2} + \left(\pi_1^{*^2} + \pi_2^*\right) \cdot \varepsilon_{t+1}$$
$$\delta_t(4) = \varepsilon_{t+4} + \pi_1^* \cdot \varepsilon_{t+3} + \left(\pi_1^{*^2} + \pi_2^*\right) \cdot \varepsilon_{t+2} + \left(\pi_1^{*^3} + 2 \cdot \pi_1^* \cdot \pi_2^* + \pi_3^*\right) \cdot \varepsilon_{t+1}$$

etc.

The error of the predictor can be written as:

$$\delta_t(h) = \varphi_0^* \cdot \varepsilon_{t+h} + \varphi_1^* \cdot \varepsilon_{t+h-1} + \ldots + \varphi_{h-1}^* \cdot \varepsilon_{t+1}. \tag{2.7}$$

If the relations (2.6) and (2.7) are written by means of the backward operator U, we obtain the connection between parameters: π_w^* and φ_w^*. Denoting:

$$\left(1 - \pi_1^* U - \ldots - \pi_{h-1}^* U^{h-1}\right) \cdot \delta_t(h) = \varepsilon_{t+h}, \tag{2.8}$$
$$\left(\varphi_0^* + \varphi_1^* U + \ldots + \varphi_{h-1}^* U^{h-1}\right) \cdot \varepsilon_{t+h} = \delta_t(h), \tag{2.9}$$

we obtain:

$$\left(1 - \pi_1^* U - \ldots - \pi_{h-1}^* U^{h-1}\right)\left(\varphi_0^* + \varphi_1^* U + \ldots + \varphi_{h-1}^* U^{h-1}\right) = 1. \tag{2.10}$$

The coefficients φ_w^* satisfy the following difference equation:

$$\varphi_w^* = \pi_1^* \cdot \varphi_{w-1}^* + \pi_2^* \cdot \varphi_{w-2}^* + \ldots + \pi_w^* \cdot \varphi_0^*, \tag{2.11}$$

$w = 1, 2, \ldots, h-1$; $\varphi_0^* = 1$.

The variance function of the forecast error of regular component is as follows:

$$\operatorname{var}[\delta_t(h)] = \sigma_\varepsilon^2 [1 + (\varphi_1^*)^2 + (\varphi_2^*)^2 + \ldots + (\varphi_{h-1}^*)^2]. \tag{2.12}$$

The general form of predictor of ARIMA(p,k,q) process is:

$$\hat{x}_{t+h} = \mathrm{f}_\mathrm{h}^* X_t + \sum_{r=1}^{p+k-1} \sum_{s=1}^{p+k-r} \mathrm{f}_{\text{h-s}}^* \beta_{s+r}^* X_{t-r} + \sum_{s=0}^{\mathrm{q-1}} \sum_{\mathrm{r}=1}^{\mathrm{q}} \mathrm{f}_{\text{h-r}}^* \gamma_{r+s} \varepsilon_{t-s}, \tag{2.13}$$

where the relation between parameters f_i^* and β_s^* is as it follows:

$$f_i^* = \beta_1^* f_{i-1}^* + \beta_2^* f_{i-2}^* + \ldots + \beta_i^* f_0^*, \quad i = 1, 2, \ldots \quad f_0^* = 1, \; f_{-k}^* = 0, \quad \beta_i^* = 0 \text{ for } i > p+k. \tag{2.14}$$

The forecast error of the ARIMA(p,k,q) process can be rewritten as:

$$\delta_t(h) = \sum_{w=0}^{h-1} \varphi_{\mathrm{w}}^{*} \varepsilon_{t+h-w} , \qquad (2.15)$$

where:

$$\varphi_w^* = \sum_{r=0}^{w} f_{w\text{-}r}^{*} \gamma_r ,$$

$$\gamma_r = 0 \quad \text{dla} \quad r > q, \qquad (2.16)$$

$$f_{-k}^{*} = 0.$$

For example, for the ARIMA(0,1,q) process we obtain:

$$\varphi_w^* = (\beta_1^* + \gamma_1) \cdot \beta_1^{*(w-1)} \qquad (2.17)$$

and the variance of the forecast error is as follows:

$$\mathrm{var}[\delta_t(h)] = \sigma_\varepsilon^2 \left[1 + \sum_{w=1}^{h\text{-}1} \left(\beta_1^* + \gamma_1\right)^2 \cdot \beta_1^{*2(w-1)} \right]. \qquad (2.18)$$

According to the (2.18) expression, the longer the forecast period is, the greater the variance of forecast error is.

Clements (2001) studies predictors of different non-stationary processes and their forecast failure.

3. Spectral representation of ARIMA(p,k,q) process

Priestley (1981) shows the spectral representation of non-stationary processes and compares spectrum of theoretical and evolutionary processes for different values of t.

If the stochastic process X_t is an output of the linear filter with an operator L and a white noise ε_t as input:

$$X_t = L \cdot \varepsilon_t , \qquad (3.1)$$

the stochastic process is shown by the following spectral presentation:

$$X_t = \int_{-\pi}^{\pi} H(\omega) \cdot e^{i \cdot \omega \cdot t} \, dZ_\varepsilon(\omega) , \qquad (3.2)$$

$t = 0, \pm 1, \ldots,$

where:

$H(\omega)$ – function of reaction frequency also called *the transfer function*,

$\int_{-\pi}^{\pi} e^{i\cdot\omega\cdot t} dZ_\varepsilon(\omega)$ – complex form of ε_t.

Simulation linear filters unchanged in time[2] transform every stationary entrance process into the stationary output process. The transfer function can be written in the complex form:

$$H(\omega) = G(\omega)\cdot e^{-i\cdot\varphi(\omega)}, \tag{3.3}$$

where:

$$G(\omega) = |H(\omega)| = \sqrt{[\mathrm{Re}\ H(\omega)]^2 + [\mathrm{Im}\ H(\omega)]^2}\ , \tag{3.4}$$

The module $H(\omega)$, which is the amplitude characteristic, is called *the gain function*;

$$\varphi(\omega) = \arg H(\omega) = \mathrm{arc\,tan}\,\frac{\mathrm{Im}\ H(\omega)}{\mathrm{Re}\ H(\omega)}, \tag{3.5}$$

The argument $H(\omega)$ is called the *function of phase angle*. Using formulas (3.3) and (3.2) we obtain:

$$X_t = \int_{-\pi}^{\pi} G(\omega)\cdot e^{i\cdot[\omega\cdot t-\varphi(\varpi)]} dZ_\varepsilon(\omega). \tag{3.6}$$

The expression (3.6) shows, that the equivalent of entrance process with frequency ω is the output with the same frequency, but weighted in an amplitude by the factor $G(\omega)$ and shifted in a phase by $\phi(\omega)$. The filtration of stationary stochastic processes modifies the amplitude of harmonic components of the process and causes phasic movement of these components.
It is said, that the filter is completely precise if the increase function and the function of the phase angle are known.
The parameter

$$\tau(\omega) = \frac{\phi(\omega)}{\omega}, \qquad \omega \neq 0 \tag{3.7}$$

measures the movement of phase in a time unit and it is called the *postponement function*. The filter L postpones the original time of the harmonic with frequency ω with $\tau(\omega)$ time units.

The ARIMA(*p*,*k*,*q*) process can be treated as output X_t of a linear filter with operator $\Pi^*(U)$, with a white noise ε_t as an input. Spectral presentation of ARIMA(*p*,*k*,*q*) process is shown in the following formula:

---

[2] Simulation linear filters unchanged in time are: $Y_t = LX_t = \sum_{s=-p}^{q} h_s X_{t-s}$, $t = 0, \pm1, \ldots$; $p, q > 0$; h_s – real; $\sum_{s=-p}^{q} \mathrm{h}_s^2 < \infty$, which are linear: $L(\alpha(\alpha_{1t} + \beta X_{2t}) = \alpha LX_{1t} + \beta LX_{2t}$ and unchanged in time. It means that if $L\cdot X_t = Y_t$, then $L\cdot X_{t+r} = Y_{t+r}$ for every *r*. See: Talaga, Zieliński (1986), § 1.5.

$$X_t = \int_{-\pi}^{\pi} e^{i\omega t} F^*(\omega) Z_\varepsilon(\omega), \tag{3.8}$$

where the transfer function $H(\omega) = F^*(\omega)$ can be rewritten:

$$F^*(\omega) = \frac{C(\omega)}{B^*(\omega)} = \frac{\sum_{r=0}^{q} \gamma_r e^{-i\omega r}}{\sum_{s=0}^{p+k} (-\beta_s^*) e^{-i\omega s}} = \sum_{j=0}^{\infty} \varphi_j^* e^{-i\omega j}, \tag{3.9}$$

and the coefficients φ_j^* are described by means of following equation:

$$\varphi_i^* = \gamma_i + \beta_1^* \cdot \varphi_{i-1}^* + \beta_2^* \cdot \varphi_{i-2}^* + \ldots + \beta_{p+k}^* \cdot \varphi_{i-(p+k)}^*, \quad i = 1, 2, \ldots \tag{3.10}$$

$$\varphi_0^* = 1; \quad \varphi_{-k}^* = 0;$$

$$\gamma_i = 0 \text{ dla } i > q.$$

With reference to formula (1.6) the coefficients β_v^*, $v = 1, 2, \ldots, p+k$ can be shown as β_s parameters:

$$\beta_r^* = \left[\binom{k}{0} \cdot \beta_r - \binom{k}{1} \cdot \beta_{r-1} + \binom{k}{2} \cdot \beta_{r-2} - \ldots \pm \binom{k}{r-1} \cdot \beta_1 - (-1)^r \cdot \binom{k}{r}\right],$$
$$r = 1, 2, \ldots, p-1;$$

$$\beta_s^* = (-1)^{s+p} \cdot$$
$$\cdot \left[\binom{k}{s-p} \cdot \beta_p - \binom{k}{s-p+1} \cdot \beta_{p-1} + \binom{k}{s-p+2} \cdot \beta_{p-2} - \ldots \pm \binom{k}{s-1} \cdot \beta_1 - (-1)^p \cdot \binom{k}{s}\right],$$
$$s = p, p+1, \ldots, p+k-1;$$

$$\beta_{p+k}^* = (-1)^k \cdot \beta_p \tag{3.11}$$

The predictor (2.16) can be rewritten as:

$$\hat{x}_{t+h} = \sum_{w=0}^{\infty} \varphi_{h+w}^* \varepsilon_{t-w}. \tag{3.12}$$

Spectral presentation of (3.13) predictor is as follows:

$$\hat{x}_{t+h} = \int_{-\pi}^{\pi} e^{i\omega\omega} \left\{ e^{i\omega h} \left[F^*(\omega) - \mathrm{F}_h^*(\omega) \right] \right\} dZ_\varepsilon(\omega), \tag{3.13}$$

where:

$$F^*(\omega) - F_h^*(\omega) = \varphi_h^* e^{-ih\omega} + \varphi_{h+1}^* e^{-i(h+1)\omega} + \ldots \tag{3.14}$$

The transfer function of (3.13) predictor is shown in equation (3.15):

$$H_h^*(\omega) = \sum_{s=h}^{\infty} \varphi_s^* e^{-i(s-h)\omega} \tag{3.15}$$

and a function of movement of ARIMA(p,k,q) process predictor is shown as follows:

$$\left|H_h^*(\omega)\right|^2 = \left[\sum_{s=h}^{\infty} \varphi_s^* \cos(s-h)\omega\right]^2 + \left[\sum_{s=h+1}^{\infty} \varphi_s^* \sin(s-h)\omega\right]^2 . \tag{3.16}$$

The gain function $G(\omega)$ has been compared with the function of the phase angle $\phi(\omega)$ of an ARIMA(1,1,1)-th ordered autoregressive-integrated moving average process and its predictor for different values of β and γ parameters.

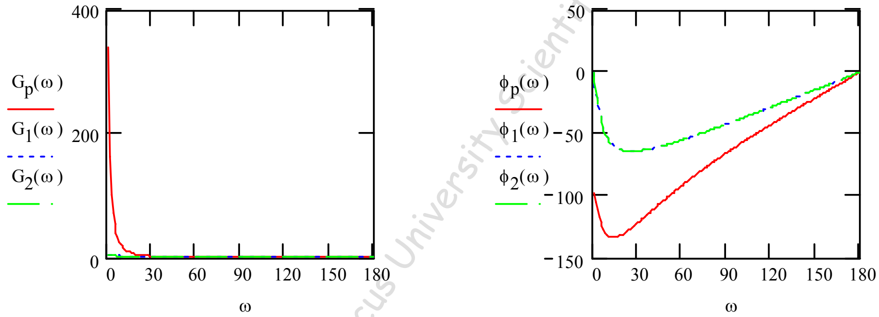

Fig. 1. The gain function of the $G_p(\omega)$ process and of the predictor of the ARIMA(1,1,1) process for h = 1 (G_1) and h = 2 (G_2); the function of the phase angle of the $\phi_p(\omega)$ process and of the predictor of ARIMA(1,1,1) process for $h = 1$ (ϕ_1) and $h = 2$ (ϕ_2), $\beta_1 = 0.9$; $\gamma_1 = -0.4$.

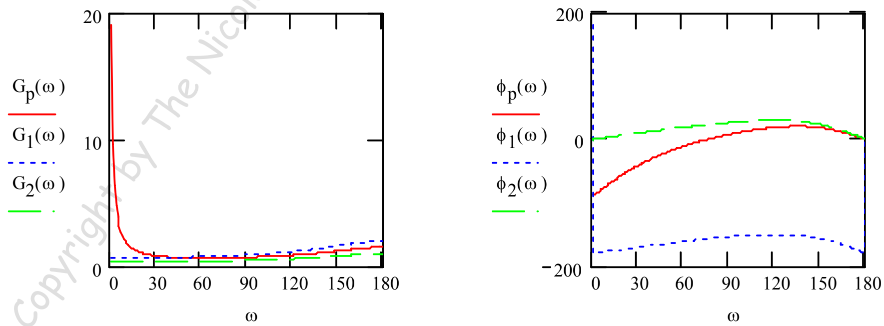

Fig. 2. The gain function of the $G_p(\omega)$ process and of the predictor of the ARIMA(1,1,1) process for h = 1 (G_1) and h = 2 (G_2); the function of the phase angle of the $\phi_p(\omega)$ process and of the predictor of ARIMA(1,1,1) process for $h = 1$ (ϕ_1) and $h = 2$ (ϕ_2), $\beta_1 = –0,5$; $\gamma_1 = –0.5$.

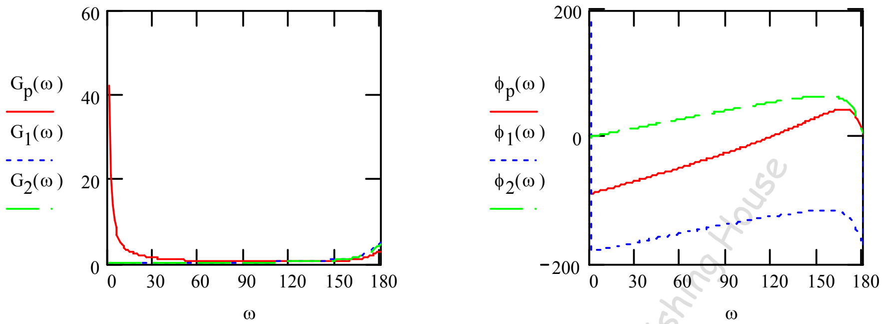

Fig. 3. The gain function of the $G_p(\omega)$ process and of the predictor of the ARIMA(1,1,1) process for h = 1 (G_1) and h = 2 (G_2); the function of the phase angle of the $\phi_p(\omega)$ process and of the predictor of ARIMA(1,1,1) process for h = 1 (ϕ_1) and h = 2 (ϕ_2), β_1 = –0,9; γ_1 = 0.4.

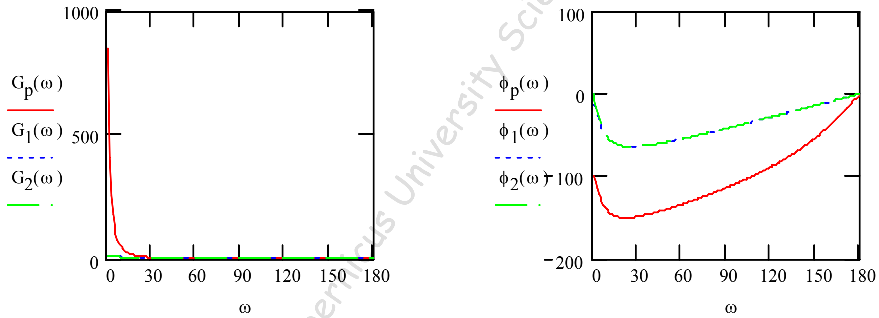

Fig. 4. The gain function of the $G_p(\omega)$ process and of the predictor of the ARIMA(1,1,1) process for h = 1 (G_1) and h = 2 (G_2); the function of the phase angle of the $\phi_p(\omega)$ process and of the predictor of ARIMA(1,1,1) process for h = 1 (ϕ_1) and h = 2 (ϕ_2), β_1 = 0,9; γ_1 = 0.5.

The values of the *gain function* $G_p(\omega)$ of the ARIMA(1,1,1) *process* show, that the filter of this process leaks highly the components of low frequencies for various values of β_1 and γ_1 parameters. The filter damps almost completely the componets of others frequencies. If $\beta_1 > 0$, the filter damps completely the components of others frequencies. If $\beta_1 < 0$, the filter leaks the components of high frequencies.

The gain function $G_h(\omega)$ of the *predictor of the* ARIMA(1,1,1) process damps the components of low frequencies and for higher frequencies it proceeds like the $G_p(\omega)$ gain function of the process. The gain function of predictors for different forecasts (h = 1 i h = 2) has almost an identical outcome. The comparison between the gain function of the process and the gain function of predictor of the ARIMA(1,1,1) process brings us to the conclusion, that the effectiveness of the predictors is greater for the components of those frequencies, for which the $G_p(\omega)$ gain function of the process and $G_h(\omega)$ gain function of the predictor are the same shape. For the ARIMA(1,1,1) process the effectiveness of the predictors is greater for short-term fluctuations.

The sign of the β_1 parameter of autoregression affects the shape of the *function of the phase angle of the process* $\phi_p(\omega)$ and the *function of the phase angle of the predictor* $\phi_h(\omega)$ of the ARIMA(1,1,1) process.

If $\beta_1 > 0$ the phase shift of the *process* is negative. In the interval of low frequencies the shift increases, and in other interval – it decreases. That means that the components of low frequencies are more lagged than the components of high frequencies. In the interval of higher frequencies the gradient of the function of the phase angle is almost constant, so the time-delay ($\phi(\omega)/\omega$) is almost equal. The negative phase shows that the output process is lagged behind the input process. The phase shift of the *predictor* is the similar shape but with less values (especially in the interval of low frequencies). The phase shift of the predictors for the forecasts: $h = 1$ and $h = 2$ is identical. The comparison between the function of the phase angle of the process and of the predictor brings us to the conclusion, that the effectiveness of the predictors is greater for the components of the high frequencies than for the low frequencies components.

If $\beta_1 < 0$, the shapes of the function of the phase angle of the process and of the predictors of ARIMA(1,1,1) process are completely different. Their signs, their values of the shift and their forecasts ($h = 1$ i $h = 2$) are different.

The considerations above say that if the shapes of the gain function and of the function of the phase angle of the process and of the predictors of the ARIMA(p,k,q) process have identical outcomes – the effectiveness of the predictors is much greater than when the shapes of those functions are completely different.

References

Box, G. E. P., Jenkins, G. M. (1983), *Analiza szeregów czasowych. Prognozowanie i sterowanie (Time Series Analysis. Forecasting and Control)*, PWN, Warszawa.

Clements, M. P. (2001), *Forecasting Non-stationary Economic Time Series*, The MIT Press, London.

Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Academic Press, London.

Talaga, L., Zieliński, Z. (1986), *Analiza spektralna w modelowaniu ekonometrycznym (Spectral Analysis in Econometric Modelling)*, PWN, Warszawa.