

Jan Purczyński*

Uniwersytet Szczeciński

WYBRANE ASPEKTY PROGNOZOWANIA Z WYKORZYSTANIEM KLASYCZNYCH MODELI TRENDU

STRESZCZENIE

W artykule rozpatrzono wybrane aspekty prognozowania z wykorzystaniem klasycznych modeli trendu. Zwrócono uwagę na możliwość wykonania błędnej prognozy dla trendu wielomianowego i potęgowego. W przypadku trendu wielomianowego niebezpieczeństwo wynika z faktu, że prognoza wyróżniająca się bardzo małym błędem *ex ante* może być obciążona bardzo dużym błędem *ex post*. Trend potęgowy niesie natomiast zagrożenia polegające na niedopasowaniu krzywej teoretycznej do danych empirycznych, co spowodowane jest wystymowaniem wartości wykładnika potęgi (wzór (6)) $b < 1$ w miejsce poprawnej wartości $b > 1$.

Słowa kluczowe: predykcja z wykorzystaniem klasycznych modeli trendu, trend wielomianowy, trend potęgowy.

1. Uwagi dotyczące wykorzystania trendu wielomianowego

Klasyczny model trendu opisany szeregiem czasowym:

$$y_t = f(t) + \varepsilon_t \quad (1)$$

* Adres e-mail: jan.purczynski@wzieu.pl.

gdzie:

- y_t – obserwacje,
- $t = 1, 2, \dots, n$,
- ε_t – składnik losowy,
- $f(t)$ – postać analityczna trendu,

są często wykorzystywane w procesie prognostycznym ze względu na ich zalety: proste wzory określające oszacowania parametrów trendu i wartość błędu *ex ante* prognozy, brak konieczności wykonania prognozy zmiennej objaśniającej w okresie prognozy (czas t). Przykładem klasycznego modelu tendencji rozwojowej jest trend wielomianowy o postaci:

$$y_t = \sum_{j=0}^N a_j t^j + \varepsilon_t \quad (2)$$

gdzie y_t, ε_t – jak we wzorze (1).

W zależności od wartości stopnia wielomianu N (wzór (2)) wyróżnia się $N = 1$ trend liniowy, $N = 2$ trend paraboliczny, $N = 3$ trend wielomianowy trzeciego stopnia, $N = 4$ trend wielomianowy czwartego stopnia.

Analizując literaturę z zakresu wykorzystania klasycznych modeli trendu w procesie prognozowania, można zauważyć dużą polaryzację poglądów na ocenę przydatności trendu wielomianowego w zagadnieniach predykcji. Najbardziej entuzjastyczną opinię można znaleźć w pracy [Milo, 2002, s. 42], z której pochodzi następujący cytat: „Użyteczność prognostyczna wielomianowych funkcji trendu była i jest oczywista zarówno dla ekonomistów [Lange, 1976; Tintner, 1952], jak i inżynierów [Johnson, Leone, 1977]. Modele wielomianowe mają niewątpliwie ważną zaletę w postaci dużej elastyczności doboru stopnia wielomianu, a tym samym stopnia przybliżenia modelu do danych. Ekonomisci najchętniej posługują się parametrycznymi funkcjami wielomianów względem zmiennej czasowej t . Wielomiany takie są naiwnymi modelami prognozowanego procesu w sensie kazu- alnym, ale ich użyteczność w prognozowaniu krótkookresowym *ex ante* jest często wyjątkowo duża”. Bardziej powściągliwe stanowisko prezentowane jest w pracy [Hozer, 1994, s. 354], gdzie stwierdza się: „W praktyce stosuje się krzywe trendu

stopnia co najwyżej trzeciego”. Podobne stanowisko reprezentują autorzy pracy [Zeliaś, Pawełek, Wanat, 2003, s. 80, 81], w której zamieszczono wzory dotyczące estymacji parametrów trendu parabolicznego i wielomianu trzeciego stopnia.

Kolejne ograniczenie dotyczące stopnia wielomianu można znaleźć w pracy [Cieślak, 2001, s. 76], gdzie spośród trendów wielomianowych wymienia się tylko trend paraboliczny.

Najbardziej radykalne stanowisko zajmuje Z. Pawłowski, który stwierdza [Pawłowski, 1973, s. 172, 173]: „Innymi słowy, nawet bardzo dobra zgodność obserwacji z trendem wielomianowym w próbie nie pozwala automatycznie zakładać, że będzie tak i w przyszłości. Zauważmy, że gdy próba liczy n obserwacji, zastosowanie wielomianu stopnia $n - 1$ zawierającego n parametrów powoduje, że wszystkie obserwacje leżą na tym wielomianie i mamy idealną zgodność danych empirycznych z modelem. Uzyskanie takiej «zgodności» nie znaczy jednak, że następną obserwacja nie odchyli się już – i to nawet znacznie – od $f(t)$ ”. Na kolejnej stronie [Pawłowski, 1973, s. 173] Z. Pawłowski konkluduje „Krytyka koncepcji trendów wielomianowych nie odnosi się jedynie do przypadku najprostszego, a mianowicie trendu liniowego”.

Z tego pobieżnego przeglądu literatury wynika pełna gama stanowisk – od wielomianu stopnia czwartego (a może i wyższego stopnia) do wielomianu stopnia pierwszego. W celu zweryfikowania tych rozbieżnych opinii rozpatrzono przykład, zaczerpnięty z pracy [Czyżycki, Hundert, Klóska, 2006, s. 67, 68], dotyczący prognozy zużycia wody z wodociągów w gospodarstwach domowych w Polsce. W cytowanej pracy prognozę wykonano z wykorzystaniem trendu hiperbolicznego.

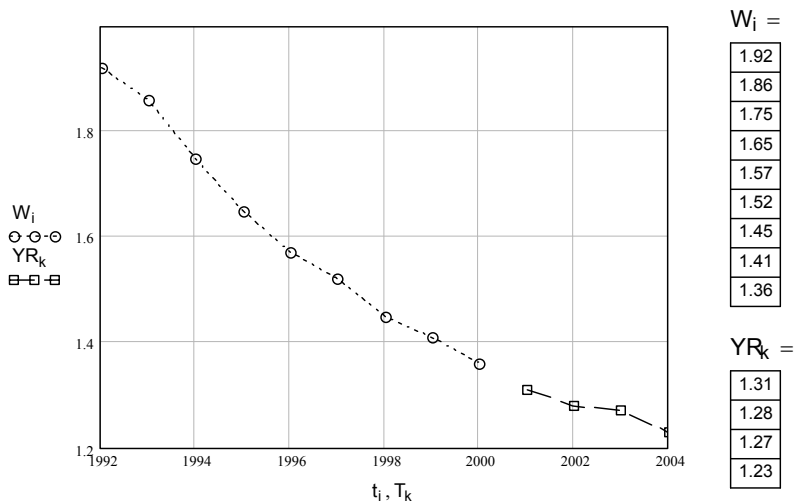
Błąd względny średniokwadratowy prognozy *ex post* wyznaczano ze wzoru:

$$BS = \frac{\sqrt{\frac{1}{4} \sum_{t_k} (YR_k - YP_k)^2}}{YS} \cdot 100 \quad (3)$$

gdzie $YS = \frac{1}{4} \sum_{t_k} YR_k$.

Wartości błędów względnych *ex post* oraz oceny współczynników zmienności dla poszczególnych stopni wielomianu wynoszą: $BS2 = 2,00\%$, $v2 = 1,13\%$ ($N = 2$); $BS3 = 8,41\%$, $v3 = 1,16\%$ ($N = 3$); $BS4 = 39,05\%$, $v4 = 0,808\%$ ($N = 4$). Z rezultatów tych wynika niebezpieczeństwo polegające na tym, że kierując się najlepszym dopasowaniem trendu do obserwacji (wartością współczynnika zmienności), zostanie wybrany trend wielomianowy czwartego stopnia, który prowadzi do błędnej prognozy. Spośród wszystkich rozpatrywanych modeli trendu najmniejszą wartość błędu *ex post* zapewniał trend paraboliczny ($N = 2$).

Rysunek 1. Zużycie wody (w tys. hm^3) z wodociągów w gospodarstwach domowych w Polsce w latach 1992–2004

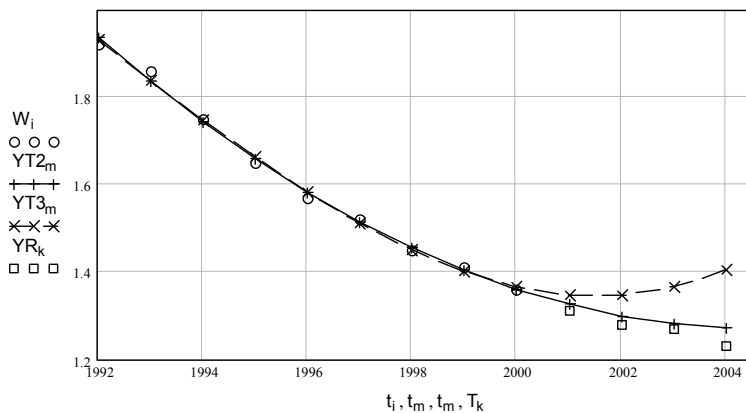


Źródło: na podstawie danych [Czyżycki, Hundert, Klóska, 2006].

Na rysunku 1 zamieszczono dane liczbowe (w tys. hm^3) zużycia wody w latach 1992–2000, oznaczone linią kropkowaną z kółkami W oraz realizację zmiennej prognozowanej w latach 2001–2004 zaznaczone prostokątami YR .

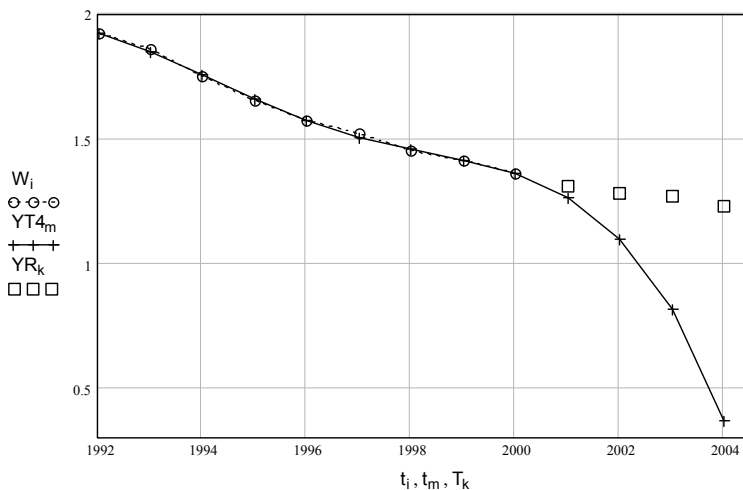
Na rysunkach 2 i 3 przedstawiano prognozy otrzymane na podstawie trendów wielomianowych: rysunek 2 – wielomian stopnia $N = 2$ i $N = 3$; rysunek 3 – wielomian stopnia $N = 4$.

Rysunek 2. Prognoza zużycia wody z wykorzystaniem trendu wielomianowego: linia przerywana z iksami $YT3$ – wielomian trzeciego stopnia, linia ciągła z plusami $YT2$ – wielomian paraboliczny ($N = 2$). Obserwacje W (w tys. hm^3) zaznaczono kółkami, a realizacje zmiennej prognozowanej YR (w tys. hm^3) w latach 2001-2004 prostokątami



Źródło: opracowanie własne.

Rysunek 3. Prognoza zużycia wody z wykorzystaniem trendu wielomianowego czwartego stopnia – linia ciągła z plusami $YT4$. Obserwacje W (w tys. hm^3) zaznaczono kółkami, a realizacje YR (w tys. hm^3) prostokątami



Źródło: opracowanie własne.

Z rysunków 2 i 3 wynika, że horyzont prognozy nie powinien przekraczać jednego roku. Więcej przykładów ilustrujących stosowanie trendu wielomianowego zamieszczono w pracy [Purczyński, 2010, s. 255–267].

Nawiązując do zaprezentowanych opinii na temat wykorzystania trendu wielomianowego, należałoby przyznać rację Z. Pawłowskiemu [Pawłowski, 1973], który ostrzegął, że uzyskanie dużej zgodności trendu wielomianowego z obserwacjami nie daje żadnej gwarancji co do jego zachowań w okresie prognozowanym. Jest to szczególnie widoczne na rysunku 3. Bardzo dobra zgodność trendu z obserwacjami została „wymuszona” przez MNK, która minimalizuje sumę kwadratów błędów. Narzucone więzy przestają obowiązywać w okresie prognozowanym, a trend przyjmuje wartości wynikające z mechanizmu rozwojowego wcześniej oszacowanych współczynników wielomianu α_j (wzór (2)). Uwzględniając, że obserwacje zawierają składnik losowy, dobre dopasowanie trendu wielomianowego należy uznać za jego wadę. Mianowicie, prowadzi to do małych wartości odchylenia standardowego reszt, a w konsekwencji do małej wartości błędu *ex ante*, co zostało podkreślone w cytowanej pracy [Milo, 2002, s. 42]¹. Powstaje zatem paradoks: małej wartości błędu prognozy *ex ante* mogą towarzyszyć duże wartości błędu prognozy *ex post*.

Wątpliwości związane ze stosowaniem trendów wielomianowych można wyjaśnić przez odwołanie się do definicji trendu zamieszczonej w pracy [Zeliaś, Pawełek, Wanat, 2003, s. 71]: „Trend jest to długookresowa skłonność do jednokierunkowych zmian (wzrostu lub spadku) wartości badanej zmiennej”. W świetle tej definicji do modelowania trendu można wykorzystać funkcję liniową oraz paraboliczną, natomiast wielomiany trzeciego i czwartego stopnia nie spełniają warunku jednokierunkowych zmian.

Podsumowując dotychczasowe rozważania, należy uznać za zasadne stosowanie trendu liniowego i parabolicznego, przyznając rację autorom prac [Cieślak, 2001, s. 76; Nowak, 1998, s. 50].

2. Trend potęgowy

Kolejnym modelem szeregu czasowego obarczonym ryzykiem błędnej prognozy jest model potęgowy. Dla zilustrowania tej tezy zaczerpnięto przykład z pracy T. Stanisza [Stanisz, 1986, s. 224–225]. Na rysunku 4 zamieszczono dane

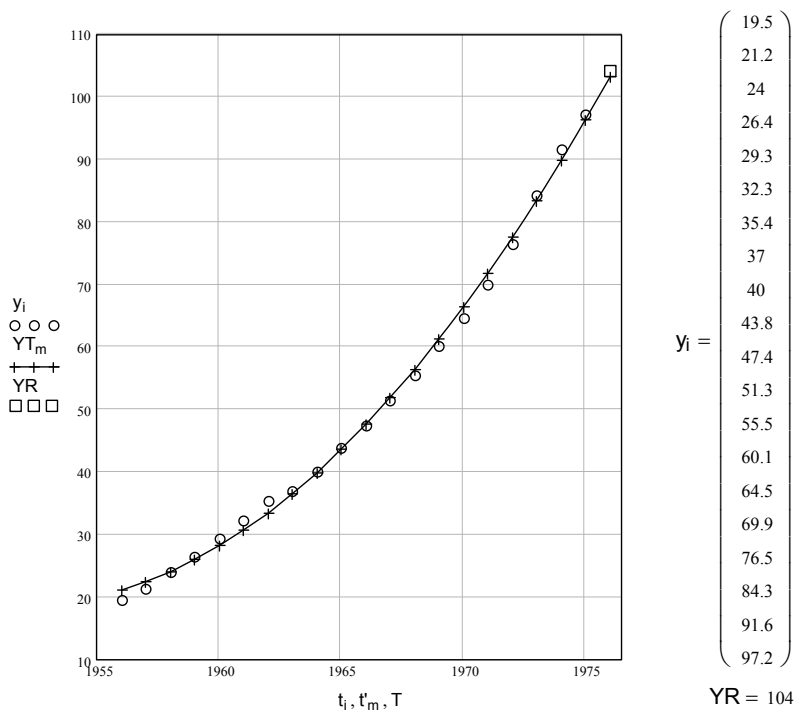
¹ „Użyteczność w prognozowaniu krótkookresowym *ex ante* jest często wyjątkowo duża”.

dotyczące produkcji energii elektrycznej (w mld KWh) w Polsce w latach 1956–1976. Wartości dotyczące 1976 roku ($YR = 104$) posłużyły do wyznaczenia błędu prognozy *ex post*.

Najpierw zastosowano trend paraboliczny, uzyskując:

$$YT = 20,037 + 0,888 \cdot t + 0,146 \cdot t^2 \tag{4}$$

Rysunek 4. Prognoza produkcji energii elektrycznej z wykorzystaniem trendu parabolicznego linia ciągła z plusami YT oznacza wartości teoretyczne. Obserwacje y (w mld KWh) zaznaczono kółkami, a realizację YR w 1976 roku prostokątem



Źródło: opracowanie własne.

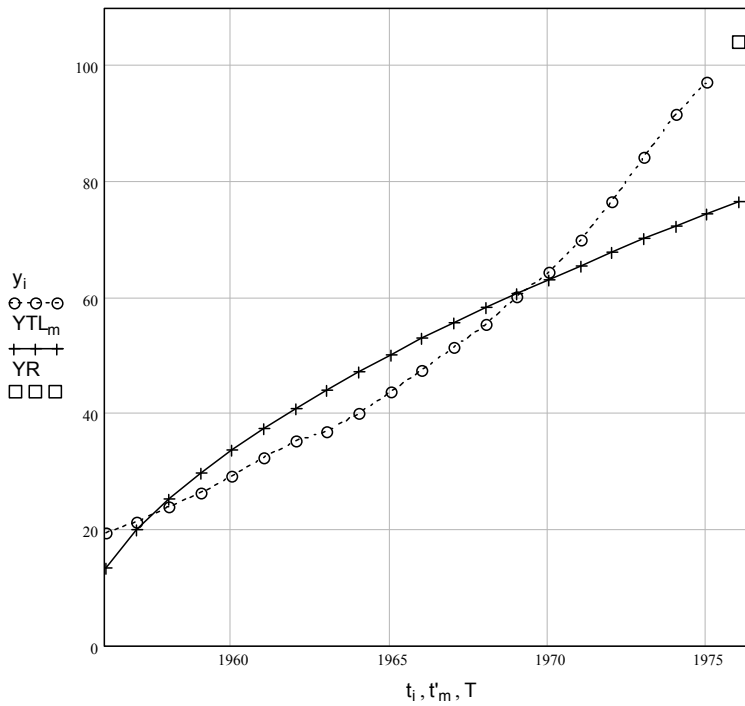
Dla tego trendu uzyskano względny błąd prognozy *ex ante* $VA = 1,51\%$ oraz względny błąd prognozy *ex post* $VP = 0,75\%$.

W pracy [Stanisz, 1986] wykorzystano trend potęgowy. Stosując metodę transformacji logarytmicznej, uzyskuje się następującą zależność:

$$YTL = 13,445 \cdot t^{0,571} \quad (5)$$

Na rysunku 5 linią ciągłą z plusami YTL zaznaczono wartości teoretyczne, a linią przerywaną z kółkami obserwacje y (mld KWh). Otrzymana na podstawie tego trendu prognoza jest obciążona błędem względnym *ex ante* $VA = 17,58\%$ oraz błędem względnym prognozy *ex post* $VP = 26,37\%$.

Rysunek 5. Prognoza produkcji energii elektrycznej z wykorzystaniem trendu potęgowego. Prostokątem zaznaczono realizację YR w 1976 roku



Źródło: opracowanie własne.

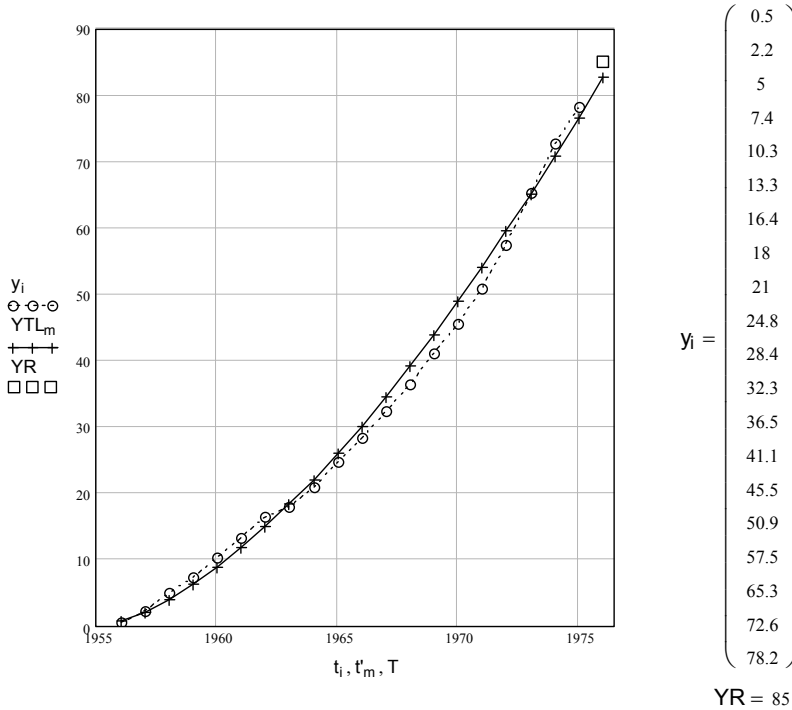
W cytowanej pracy [Stanisz, 1986] podano parametry trendu, jednak nie wykonano rysunku, w wyniku czego zamieszczony tam przykład ilustrował zastosowanie trendu potęgowego.

Zgodnie z rysunkiem 4, dane empiryczne wyznaczają krzywą (linia kropkowana) będącą funkcją wypukłą, co oznacza, że wykładnik potęgi b we wzorze (6):

$$YT = a \cdot t^b \quad (6)$$

powinien przyjąć wartość większą od 1 ($b > 1$). Wyestymowana wartość $\hat{b} = 0,571$ ($\hat{b} < 1$) odpowiada natomiast funkcji wklęsłej – linia ciągła na rysunku 5. Autor artykułu wielokrotnie spotkał się z tym paradoksem, przy czym problem pojawiał się wówczas, gdy dane empiryczne wyznaczały funkcję wypukłą tak jak na rysunku 5. Gdy dane wyznaczały funkcję wklęsłą (wypukłością do góry), nie występowało opisane zjawisko. Przyczyny należy upatrywać w zbyt dużej wartości obserwacji dla 1956 roku, co odpowiadało $t = 1$ we wzorze (6). Trend potęgowy (wzór (6)) dla $t = 0$ przyjmuje wartość $YT = 0$. Oznacza to, że wartości $t = 1$ nie może odpowiadać zbyt duża liczba. Uwzględniając dane zamieszczone na rysunku 4, od wartości obserwacji (oraz realizacji YR) odjęto stałą $A = 19$. Wyniki obliczeń przedstawiono na rysunku 6, z którego wynika stosunkowo duża zgodność krzywej teoretycznej (YTL) z danymi empirycznymi (y).

Rysunek 6. Prognoza produkcji energii elektrycznej z wykorzystaniem trendu potęgowego – linia ciągła z plusami YTL . Prognozę wykonano dla zmodyfikowanych wartości obserwacji (odjęto liczbę $A = 19$)



Źródło: opracowanie własne.

Wartość prognozy $YP = 82,684$ obarczona jest błędem względnym *ex post* wynoszącym $VP = 2,72\%$. Jeżeli powrócić do danych pierwotnych (dodając stałą A), to wynik prognozy wynosi $YP = 101,684$, a wartość błędu *ex post* maleje do $VP = 2,23\%$. Wynika stąd następujący algorytm:

- od wartości obserwacji y odjąć stałą A tak dobraną, aby dla $t = 1$ uzyskać wartość bliską zeru,
- wykonać prognozę dla zmienionych wartości y ,
- do wyniku prognozy dodać stałą A .

Podsumowanie

W artykule zwrócono uwagę na możliwość wystąpienia dużych wartości błędu predykcji dla wybranych modeli trendu. W przypadku trendu wielomianowego niebezpieczeństwo wynika z faktu, że prognoza wyróżniająca się bardzo małym błędem *ex ante* może być obciążona bardzo dużym błędem *ex post*. Oznacza to, że prognozy budowane na podstawie predyktorów opartych na modelach trendu z wielomianami wyższych stopni nie mogą zbyt daleko wybiegać w przyszłość. Trend potęgowy niesie natomiast zagrożenia polegające na niedopasowaniu krzywej teoretycznej do danych empirycznych, co spowodowane jest wyestymowaniem wartości wykładnika potęgi (wzór (6)) $b < 1$ w miejsce poprawnej wartości $b > 1$.

Literatura

- Cieślak M. (red.) (2001), *Prognozowanie gospodarcze. Metody i zastosowanie*, Wydawnictwo Naukowe PWN, Warszawa.
- Czyżycki R., Hundert M., Klóska R. (2006), *Wybrane zagadnienia z prognozowania*, Wydawnictwo Economicus, Szczecin.
- Hozer J. (red) (1994), *Statystyka, cz. II. Wnioskowanie statystyczne*, Uniwersytet Szczeciński, Szczecin.
- Johnson N., Leone F. (1977), *Statistics and Experimental Design*, Vol. 1, Wiley, New York.
- Lange O. (1976), *Wstęp do ekonometrii*, w: *Dzieła*, t. 5, PWE, Warszawa.
- Milo W. (red) (2002), *Prognozowanie i symulacja*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Nowak E. (red) (1998), *Prognozowanie gospodarcze. Metody, modele, zastosowania, przykłady*, Agencja Wydawnicza Placet, Warszawa.
- Pawłowski Z. (1973), *Prognozowanie gospodarcze*, PWN, Warszawa.
- Purczyński J. (2010), *Wybrane problemy stosowania trendu wielomianowego w prognozowaniu gospodarczym*, *Ekonomiczne Problemy Usług* nr 60.
- Stanisz T. (1986), *Funkcje jednej zmiennej w badaniach ekonomicznych*, PWN, Warszawa.
- Tintner G. (1952), *Econometrics*, Wiley, New York.
- Zeliaś A., Pawełek B., Wanat S. (2003), *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, Wydawnictwa Naukowe PWN, Warszawa.

SELECTED ASPECTS OF FORECASTING WITH THE APPLICATION OF CLASSICAL TREND MODELS

Abstract

In this paper selected aspects of forecasting with the application of classical trend models are examined. Furthermore the possibility of making a wrong forecast for polynomial and power trends is pointed out. In the case of the polynomial trend, the risk results from the fact that the forecast exhibiting a very small value of *ex ante* may be burdened with a very large value of *ex post*. As far as the power trend is considered, there is the risk that the theoretical distribution will not fit empirical data, which is caused by estimating the value of the index (power) (equation (6)) $b < 1$ instead of the correct value $b > 1$.

Translated by Ewa Stefanowska

Keywords: predicting with the application of classical trend models, polynomial trend, power trend.

Kod JEL: C02, C51, C53, C63.