

*Agnieszka Tłuczak, Mirosława Szewczyk*

**EFEKTYWNOŚĆ MODELI  
AUTOREGRESYJNYCH  
W PROGNOZOWANIU CEN PRODUKTÓW  
ROLNYCH W POLSCE**

**Słowa kluczowe:** *modele autoregresyjne, prognozowanie, ceny produktów rolnych*

**Abstrakt:** Prognozowanie cen produktów rolnych odgrywa dużą rolę we wspomaganiu decyzji produkcyjnych w gospodarstwach rolnych. Poprawne wyznaczenie prognoz cen produktów rolnych pozwala ograniczyć ryzyko związane z prowadzeniem działalności gospodarczej. W opracowaniu autorki przedstawily możliwość zastosowania modeli autoregresyjnych, za pomocą których wyznaczono prognozy cen podstawowych produktów rolnych w skupie na drugie półrocze 2010 roku.

# THE EFFECTIVENESS OF THE AUTOREGRESSIVE MODELS IN FORECASTING THE AGRICULTURAL PRICES IN POLAND

**Keywords:** *autoregressive models, forecasting, agricultural prices*

**Abstract:** The forecast of agricultural prices is one of the most important factors in making decision on production farms. The appropriate forecast allows for limiting the risk connected with one's economic activity. In this study autoregressive models have been used, which helped to determine the price forecast for agricultural products in the purchasing centers in the second half of 2010. To determine the quality of forecast the average ex-post errors of the past forecasts have been used. The achieved results show that autoregressive models are an effective tool in forecasting the agricultural prices in Poland.

## WPROWADZENIE

We współczesnym, dynamicznie rozwijającym się, świecie coraz większego znaczenia nabiera umiejętność prognozowania zjawisk społeczno-gospodarczych. Kluczową umiejętnością współczesnych menedżerów jest zdolność do wykorzystania dostępnych narzędzi w celu sporządzania prognoz. Wiedza związana z wyznaczaniem prognoz jest coraz bardziej pożądana w procesie zarządzania zjawiskami ekonomicznymi, kiedy to rezultat decyzji podejmowanych dzisiaj jest w dużym stopniu uzależniony od tego jak sytuacja będzie kształtować się w przyszłości. Wyznaczanie prognoz prowadzi do zmniejszenia niepewności oraz przyczynia się do zwiększenia trafności podejmowanych decyzji, a tym samym do eliminacji strat w różnych sferach gospodarczych (Sobczyk 2008). W rolnictwie, podobnie jak w innych rodzajach działalności gospodarczej, istnieje potrzeba posiadania informacji o przyszłym stanie określonych zjawisk. Informacje te są konieczne do podejmowania trafnych decyzji, zarówno w gospodarstwach hodujących zwierzęta przeznaczonych na rzeź czy zajmujących się uprawą zbóż, zakładach przerabiających je, czy jednostkach handlujących przetworzonymi produktami rolnymi (Majewski 2006). W literaturze można zapoznać się z wieloma metodami prognozowania cen produktów rolnych. W dotychczasowych badaniach autorki wykazują, iż efektywnym narzędziem prognozowania cen produktów rolnych są między innymi modele

adaptacyjne czy analiza harmoniczna (Tłuczak 2008a; 2008b; 2009). W niniejszej pracy zaprezentowano możliwość wykorzystania szeroko prezentowanych w literaturze modeli autoregresyjnych do prognozowania cen wybranych produktów rolnych w skupie.

## **SYTUACJA NA RYNKU ROLNYM W POLSCE**

W produkcji rolnej ważnym czynnikiem powodującym zmiany wielu zmiennych są warunki klimatyczne. Niekorzystne warunki pogodowe w okresie wegetacji (chłodne noce, nadmiar opadów, które powodują często podtopienia i powodzie, z drugiej strony wysokie temperatury, czy niedostateczna ilość opadów) oraz trudności z wykonaniem niektórych zabiegów agrotechnicznych nierzadko wpływają na obniżenie plonów. Wielkość plonów z kolei wpływa na poziom cen, jakie producenci rolni mogą uzyskać w punktach skupu.

W okresie styczeń 1996 - maj 2010 ceny pszenicy kształtowały się w granicach od 34 do 92 PLN za 100 kg, średnia cena tego zboża kształtowała się na poziomie 52 PLN/100 kg. Wartość współczynnika zmienności, który wynosi 22%, wskazuje na średnie zróżnicowanie cen na rynku pszenicy, natomiast współczynnik korelacji pomiędzy ceną pszenicy w badanym okresie a wielkością skupu tego produktu wyniósł  $-0,24$ , co oznacza, że wraz ze wzrostem podaży tego ziarna na rynku jego cena maleje. Sytuacja na tym rynku w kolejnych latach kształtowała się odmiennie. Cały badany okres można podzielić na 6 podokresów, z których każdy charakteryzował się odmienną tendencją. Średnioroczne tempo zmian – w tym przypadku wzrostu cen – wyniosło zaledwie 0,34%. Zmiany sytuacji na rynku pszenicy miały miejsce w sierpniu 1996 roku, marcu 1999 roku, marcu 2001 roku, kwietniu 2003 roku, lipcu 2004 roku oraz w grudniu 2005 roku. Wiązało się to ściśle z wielkością skupu tego ziarna. Wielkość skupu wzrosła z 27 tysięcy ton w lipcu 1996 do 625 tys. ton w sierpniu 1996 roku. Tak duży wzrost wielkości skupu spowodował spadek ceny o około 10 PLN/100 kg. Kolejnym momentem zwrotnym był marzec 1999 roku – wówczas ceny pszenicy w skupie zaczęły rosnąć, podaż tego zboża na rynku nieznacznie obniżyła się, a rosnąca tendencja utrzymywała się do marca 2001 roku. Po okresie spadku podaży pszenicy na rynku wzrosła, co było efektem sprzyjających warunków pogodowych. Zwiększona podaż spowodowała obniżkę cen przez kolejne kilka miesięcy. Niekorzystna sytuacja utrzymywała się do kwietnia 2003 roku, wówczas to zaczęły obowiązywać unijne regulacje rynku zbóż. Objęcie polskiego rynku mechanizmami Wspólnej Polityki Rolnej spowodowało wzrost cen o około 50%. W latach 2004 – 2006 można zaobserwować wzrost podaży pszenicy na

rynku, w 2005 roku do skupu trafiło dwa razy więcej ziarna, niż w roku 1996. Konsekwencją wzrostu podaży był spadek cen z 70 PLN /100 kg w maju 2004 roku do poziomu 36 PLN /100 kg w grudniu 2005 roku. Nieprzyjające warunki atmosferyczne (wegetacyjne) w 2006 i 2007 roku spowodowały obniżenie się ilości ziarna pszenicy dostępnego na rynku. W tym samym okresie zmniejszył się również poziom zapasów interwencyjnych w magazynach Agencji Rynku Rolnego. Sytuacja taka spowodowała wzrost cen pszenicy w latach 2005 – 2007 o 130% z poziomu 36,12 PLN /100kg w grudniu 2005 roku do 83 PLN /100 kg w listopadzie 2007 roku.

W tym samym okresie ceny żyta wahały się w granicach od 25 do 74 PLN za 100 kg, natomiast przeciętnie za 100 kg tego zboża płacono w skupie 39 PLN; przeciętne odchylenia od średniej ceny żyta wyniosły 11,24 PLN /100kg. Wartości te pozwoliły na obliczenie współczynnika zmienności, który osiągnął poziom 28%, na podstawie czego można stwierdzić że sytuacja na rynku żyta była zróżnicowana. W latach 1996 – 2010 indeks zmian cen żyta wyniósł zaledwie 0,59%, lecz sytuacja kształtowała się różnorodnie w kolejnych latach. Największy spadek cen żyta w skupie przypadł na 2004 rok, kiedy to ceny obniżyły się z 50,76 PLN /100kg do 31,49 PLN /100 kg. Natomiast największy wzrost cen żyta w skupie odnotowano pomiędzy styczniem a grudniem 2006 roku i wyniósł on 63%. Zmieniająca się sytuacja na rynku żyta pozwoliła na wyznaczenie następujących punktów zwrotnych: grudzień 1996 roku, czerwiec 1999, kwiecień 2001, marzec 2003, lipiec 2004 oraz maj 2005 roku. Na rynku żyta, tak samo jak na rynku pszenicy, występują powtarzające się cyklicznie w okresie żniw nadwyżki podaży. Zwiększona podaż jest zazwyczaj zagospodarowywana przez Agencję Rynku Rolnego, która prowadzi wówczas interwencyjny skup ziarna. Obniżki, bądź ewentualnie podwyżki cen żyta w skupie są powiązane z sytuacją podażową, która determinuje poziom cen żyta. Od 2004 roku roczna wielkość skupu żyta maleje tak, że w 2007 roku do skupu trafiło łącznie o 253 tys. ton mniej żyta, niż w 2004. Tak duży spadek podaży ziarna na rynku spowodował wzrost ceny ze średniego poziomu 42,32 PLN /100kg w 2004 roku do 60,92 PLN /100 kg w 2007. W całym badanym okresie średnioroczne tempo zmian cen żyta wyniosło - 4% (Biuletyn Statystyczny).

W analizowanym okresie, styczeń 1996 - maj 2010, ceny bydła w skupie wahały się w granicach od 2,39 do 4,91 PLN za 1 kg żywca, średnio osiągając w rozważanym okresie poziom 3,30 PLN/1 kg. Spożycie mięsa wołowego w ostatnich latach w Polsce wyraźnie maleje na korzyść mięsa wieprzowego i drobiu. Sytuacja ta jest odzwierciedleniem rosnącej ceny tego gatunku mięsa. Ceny na rynku mięsa wołowego charakteryzują się

umiarkowaną zmiennością, czego dowodzi współczynnik zmienności cen na poziomie 23%. Zauważalny jest jednak znaczny wzrost cen bydła w pierwszym roku członkostwa polski w UE – w tym okresie cena wzrosła aż o 53%. Na rynku mięsa wołowego sytuacja zmieniała się kilka razy, przy czym najważniejsze zmiany zaszły w grudniu 2000, kwietniu 2001, październiku 2003, lipcu 2004, styczniu 2005, sierpniu 2005 oraz styczniu 2006 roku. Największe zmiany można zaobserwować w 2001 oraz 2004 roku. W kwietniu 2001 roku ilość skupionego mięsa wołowego wzrosła dwukrotnie, jednak tak duży wzrost skupu nie spowodował znaczących obniżek cen. W 2004 roku, po wstąpieniu Polski do Unii Europejskiej i otwarciu polskiego rynku, polskie mięso znalazło odbiorców poza granicami kraju, a ceny mięsa wołowego w skupie pomiędzy kwietniem a czerwcem 2004 roku wzrosły o około 40%.

Sytuacja na rynku żywca wieprzowego jest ściśle związana z występującym w naturze tzw. „cyklem świńskim”<sup>1</sup>, zatem na rynku żywca wieprzowego zachodzą znaczne zmiany cen i warunków gospodarowania. Tak też cena wieprzowiny w skupie ustala się pod wpływem działania mechanizmu rynkowego i w głównej mierze zależy od popytu na mięso wieprzowe oraz jego podaży na rynku. Ceny trzody osiągały porównywalny poziom do cen bydła od 2,46 do 5,1 PLN/1 kg, a średnio kształtowały się na poziomie 3,72 PLN za 1 kg. Analizując współczynnik zmienności, który wyniósł 15% dla trzody chlewnej, można uznać iż rynek żywca wieprzowego jest względnie stabilny. Pomimo niższej wartości współczynnika zmienności cen, na rynku trzody chlewnej można zaobserwować więcej punktów zwrotnych aniżeli na rynku bydła. Na rynku trzody chlewnej zmiana tendencji nastąpiła w marcu 1998 roku, we wrześniu 1999 roku, w listopadzie 2001 roku, w maju 2003 roku oraz w październiku 2004 roku. Po wzroście cen w pierwszym kwartale 1998 roku ceny w skupie żywca wieprzowego zaczęły się stopniowo obniżać. Było to skutkiem znaczącego wzrostu wielkości skupu trzody chlewnej – z 214 tys. ton w trzecim kwartale 1997 do poziomu 324 tys. ton w drugim kwartale 1998. Kolejny punkt zwrotny na

---

<sup>1</sup> Cykl świński definiowany jest jako ciągłe ruchy produkcji i cen żywca wieprzowego wokół długookresowego trendu, powtarzające się co pewien stały okres. Pomimo trwających od wielu lat badań nad cyklicznością chowu trzody chlewnej nie stworzono jednej koncepcji, wobec czego do tej pory ścierają się dwie teorie. Jedna źródeł cyklu poszukuje w czynnikach zewnętrznych i za bezpośrednią przyczynę występowania fluktuacji uznaje się niestabilność opłacalności chowu wywołaną zmianami cen pasz i żywca wieprzowego, złą polityką rządu itp. Według drugiej, przyczyn cykliczności należy szukać w samej istocie cyklu. Cykle świńskie zalicza się w tym przypadku do tzw. cykli specjalnych (towarowych), tj. mających swój własny mechanizm, niezależny od mechanizmu zmian koniunkturalnych. Szczegółowo mechanizm cyklu świńskiego przedstawiony jest w pracach: Stepień (2009, s. 331-335), Kwaśnicki (2010).

rynku trzody chlewnej nastąpił w trzecim kwartale 1999 roku. Na zaistniałą sytuację miał wpływ utrzymujący się niski stan pogłowia trzody oraz mniejszy, niż w okresach wcześniejszych, popyt na mięso wieprzowe. Zwrot sytuacji na omawianym rynku miał miejsce w czwartym kwartale 2001 roku. Od wiosny 2000 roku ceny mięsa wieprzowego systematycznie rosły w wyniku malejącej podaży tego rodzaju mięsa. Początkowo wzrost cen doprowadził do wygaśnięcia spadkowej tendencji pogłowia trzody chlewnej, zaś późniejszy spadek do ponownego jego rozwoju. W sytuacji zwiększającej się na rynku podaży wieprzowiny oraz spadku cen Agencja Rynku Rolnego zagospodarowała ówczesne nadwyżki rynkowe oraz wprowadziła po raz pierwszy dopłaty do prywatnego przechowywania. Dość łatwym wydaje się wytłumaczenie zmian, jakie zaszły w drugim kwartale 2003 roku oraz trzecim kwartale 2004 roku. W drugim kwartale 2003 ceny większości polskich produktów rolnych wzrosły w związku z uruchomieniem programów przedakcesyjnych, mających na celu wsparcie polskich rolników. Natomiast w trzecim kwartale 2004, po nagłym wzroście cen w pierwszym i drugim kwartale 2004 roku, spowodowanym akcesją Polski do Unii Europejskiej, ceny trzody chlewnej rolnych ustabilizowały się na średnim poziomie 3,65 PLN /1kg (Biuletyn Statystyczny 2010).

## **MODELE AUTOREGRESYJNE ORAZ MIARY DOKŁADNOŚCI PROGNOZ**

Jednym z głównych zadań modelowania ekonometrycznego jest określenie prawidłowości zmian poziomu badanego zjawiska w czasie. W tym celu stosuje się modele dynamiczne, których cechą charakterystyczną jest uwzględnienie czynnika czasu. Modele takie pozwalają na uwzględnienie zachodzących w czasie zmian w relacjach między zmiennymi. W modelowaniu ekonometrycznym często stosuje się modele, w których występują opóźnione wartości zmiennych. Spowodowane jest to tym, że często zdarza się, że wpływ zmiennej lub zmiennych objaśniających na zmienną objaśnianą nie jest natychmiastowy, lecz występuje z pewnym opóźnieniem. Przyczyny występowania opóźnień czasowych w relacjach ekonomicznych można podzielić na trzy grupy: psychologiczne, technologiczne, instytucjonalno-prawne. Większość opóźnień występujących w rolnictwie związana jest z:

- przyzwyczajeniami producentów rolnych do dotychczasowego sposobu gospodarowania;
- często kosztowną zmianą kierunku produkcji;

- zobowiązań wynikających z kontraktów, które uniemożliwiają natychmiastową zmianę kierunku produkcji (Gruszczynski 2004).

Do grupy modeli w których występują zmienne opóźnione należą modele autoregresyjne, które wywodzą się z szerszej klasy modeli regresji i znajdują szerokie zastosowanie w modelowaniu procesów gospodarczych. W modelach tych wartość bieżąca zmiennej jest wyrażona jako skończona kombinacja jej wartości przeszłych. Modele autoregresyjne mogą być stosowane do modelowania stacjonarnych szeregów czasowych bądź szeregów czasowych sprowadzalnych do stacjonarnych.

Do zapisu postaci modelu autoregresyjnego rzędu  $p$  używa się zwyczajowo notacji:

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t \quad (1)$$

gdzie:

$y_t, y_{t-1}, \dots, y_{t-p}$  – wartość zmiennej prognozowanej w momencie lub okresie  $t, t-1, \dots, t-p$ ,

$\varepsilon_t$  – proces resztowy (biały szum)<sup>2</sup>,

$\varphi_0, \varphi_1, \dots, \varphi_{t-p}$  – parametry modelu,

$p$  – rząd autoregresji, maksymalne opóźnienie zmiennej objaśnianej.

Proces budowy modeli autoregresyjnych przebiega zazwyczaj w czterech etapach polegających na: sprawdzeniu, czy dany szereg czasowy jest stacjonarny<sup>3</sup> oraz na ustaleniu maksymalnego rzędu opóźnień, estymacji parametrów modelu, weryfikacji oraz prognozowaniu.

W pierwszej kolejności należy sprawdzić czy rozważany szereg czasowy jest szeregiem stacjonarnym<sup>4</sup>. Najczęściej stosowanym testem służącym

<sup>2</sup> Proces resztowy w modelu autoregresyjnym jest procesem białego szumu, co znaczy tyle, że charakteryzuje się następującymi własnościami: 1.  $E(\varepsilon_t)=0$ , 2.  $D^2(\varepsilon_t)=\sigma^2$ , 3.  $Cov(\varepsilon_t, \varepsilon_s)=0$  dla  $t \neq s$ .

<sup>3</sup> Wiele szeregów czasowych opisujących zjawiska ekonomiczne cechuje brak stacjonarności. Niestacjonarne szeregi czasowe przekształcane są do szeregów stacjonarnych poprzez obliczenie kolejnych przyrostów  $d$  razy, wówczas szereg czasowy jest zintegrowany w stopniu  $d$ :  $y_t \sim I(d)$ . Każdy szereg zintegrowany w stopniu co najmniej pierwszym jest szeregiem niestacjonarnym. Obliczanie przyrostów nosi nazwę różnicowania szeregu, szereg  $I(d)$  staje się stacjonarny po  $d$ -krotnym zróżnicowaniu. Różnicując szeregi czasowe należy pamiętać, że teorie ekonomiczne, na których oparte jest modelowanie ekonometryczne, na ogół opisują relacje długookresowe między zmiennymi przy użyciu ich poziomów. Różnicowanie szeregów czasowych może spowodować, że model przestanie opisywać zależności długookresowe między zmiennymi. W przypadku badania zależności długookresowych należy rozważyć uważnie konieczność wprowadzania przyrostów zmiennych.

<sup>4</sup> Stacjonarność szeregu czasowego może być rozumiana jako: brak trendu, brak systematycznych zmian w wariancji, brak wahań periodycznych. W badaniach ekonometrycz-

do badania stacjonarności w szerszym sensie jest test Dickey'a-Fullera (DF – Dickey – Fuller test), test ten zwany jest również testem pierwiastka jednostkowego. Równanie podlegające testowaniu przyjmuje postać:

$$y_t = \rho y_{t-1} + \varepsilon_t. \quad (2)$$

Hipoteza zerowa mówi o istnieniu co najmniej jednego pierwiastka jednostkowego w badanym szeregu czasowym, natomiast alternatywna o jego braku, co zapisujemy następująco;

$$\begin{aligned} H_0: \rho=1 &\Rightarrow y_t \sim I(1), \\ H_1: \rho < 1 &\Rightarrow y_t \sim I(0)^6. \end{aligned}$$

Praktyczne zastosowanie testu wymaga przekształcenia równania (2) do postaci:

$$\Delta y_t = \delta y_{t-1} + \varepsilon_t \quad (3)$$

Wówczas odpowiednie hipotezy przyjmują postać:

$$\begin{aligned} H_0: \delta=0 &\Rightarrow y_t \sim I(1), \\ H_1: \delta < 0 &\Rightarrow y_t \sim I(0). \end{aligned}$$

Równanie (2) może zawierać składnik deterministyczny w postaci stałej, trendu liniowego oraz składnika sezonowego. Odpowiednie modyfikacje równania (2) mają postaci:

nych wyróżnia się stacjonarność w węższym sensie oraz stacjonarność w szerszym sensie. Zweryfikowanie stacjonarności w węższym sensie przysparza wiele trudności, użyteczne staje się zatem zweryfikowanie hipotez o stacjonarności w szerszym sensie. Szereg czasowy nazywamy stacjonarnym w szerszym sensie, jeżeli średnie i wariancje są skończone i stałe w czasie, a kowariancje zależą jedynie od różnicy okresów pomiędzy dwoma obserwowanymi zmiennymi. Zagadnienie stacjonarności szeregów czasowych w węższym i szerszym sensie zostało opisane w pracy Hamilton (1994, s. 45-46), Chatfield (2004, s. 34), Gruszczyński i inni (2004, s. 181-183).

<sup>5</sup> Równanie (2) opisuje model bez stałej.

<sup>6</sup> Tak postawione hipotezy oznaczają odpowiednio:  $H_0$ : proces jest zintegrowany rzędu pierwszego,  $H_1$ : proces jest stacjonarny. Najczęściej wartości rzędu opóźnień powyżej 3 sugerują iż proces nie jest w ogóle procesem zintegrowanym. Porównaj: Osińska (2007, s. 307); Gruszczyński, Podgórska (2004, s. 187).



$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t, \quad y_t = \alpha_0 + \alpha_1 t + \rho y_{t-1} + \varepsilon_t, \quad \text{lub} \quad y_t = \sum_{k=1}^m d_k Q_{kt} + \rho y_{t-1} + \varepsilon_t.$$

Statystyka Dickey'a-Fullera sprawdzająca omówione hipotezy dana jest następującą formułą:

$$DF = \frac{\hat{\delta}}{s(\hat{\delta})} \quad (4)$$

gdzie:

$\hat{\delta}$  - ocena parametru z równania (3) oszacowanego klasyczną metodą najmniejszych kwadratów,

$s(\hat{\delta})$  - średni błąd szacunku parametru  $\delta$ .

Statystyka DF posiada niestandardowy rozkład lewostronnie asymetryczny (Osińska 2006, Osińska 2007). Tablice wartości krytycznych dla testu DF opublikowano między innymi w pracy Charemzy i Deadmana z 1997 roku<sup>7</sup>. Słabością testu DF jest to, iż nie uwzględnia on możliwości wystąpienia autokorelacji składnika losowego  $\varepsilon_t$ . Rozwiązaniem zaproponowanym przez Dickey'a-Fullera w 1981 roku jest zastosowanie rozszerzonego testu Dickey'a-Fullera (test ADF)(Charemza, Deadman 1997). Procedura przeprowadzania testu ADF jest analogiczna jak w teście DF, jednak inne są równania podlegające testowaniu:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^k \alpha_i \Delta y_{t-i} + \varepsilon_t \quad (5)$$

Hipotezy mają postaci:

$$H_0: \delta=0 \Rightarrow y_t \sim I(1),$$

$$H_1: \delta < 0 \Rightarrow y_t \sim I(0).$$

Dla rozszerzonego testu Dickey'a - Fullera (ADF - augmented Dickey-Fuller *test*) stosuje się te same tablice wartości krytycznych co dla testu Dickey'a - Fullera (test DF).

---

<sup>7</sup> Wartości krytyczne testu są różne dla modeli ze stałą i bez niej oraz modeli ze zmianą czasową i bez tej zmiennej. Weryfikacja hipotezy o niestacjonarności szeregów w każdym z tych modeli wymaga zastosowania innych tablic, jednak sposób liczenia statystyki i interpretacja testu pozostaje bez zmian (Porównaj Charemza, Deadman 1997, s. 255 – 260).

Kolejny etap polega na ustaleniu rzędu autoregresji, w tym celu wykorzystuje się funkcję autokorelacji (ACF - *autocorrelations function*) oraz autokorelacji cząstkowej (PACF - *partial autocorrelations function*).

Funkcja ACF wskazuje zależność pomiędzy odległymi w czasie obserwacjami, natomiast funkcja PACF wskazuje czysty związek między obserwacjami<sup>8</sup>.

W praktycznych zastosowaniach, w celu jednoznacznego ustalenia rzędu opóźnień, analizę funkcji autokorelacji oraz autokorelacji cząstkowej uzupełnia się testem Quenouille'a, w którym pod rozwagę bierze się model autoregresyjny (1). W teście tym bada się istotność współczynnika autokorelacji cząstkowej a hipotezy mają postać:

$$\begin{aligned} H_0: \varphi_{pp} &= 0, \\ H_1: \varphi_{pp} &\neq 0, \end{aligned}$$

gdzie  $\varphi_{pp}$  jest współczynnikiem autokorelacji cząstkowej.

Sprawdzianem w teście Quenouille'a jest statystyka postaci:

$$t = \frac{\hat{\varphi}_{pp}}{s(\hat{\varphi}_{pp})}, \quad (6)$$

gdzie:  $\hat{\varphi}_{pp}$  - ocena współczynnika autokorelacji cząstkowej,

$s(\hat{\varphi}_{pp})$  - średni błąd szacunku współczynnika autokorelacji cząstkowej.

Statystyka ta ma rozkład t-Studenta, jeżeli  $t \geq 2$  to współczynnik jest statystycznie istotny. Przyjmuje się, że błąd standardowy współczynnika autokorelacji cząstkowej wynosi  $s(\hat{\varphi}_{pp}) = \frac{1}{\sqrt{n}}$  (Osińska 2007, Hatanaka 2003)<sup>9</sup>.

W kolejności dokonuje się estymacji parametrów modelu autoregresyjnego. Można tego dokonać za pomocą klasycznej metody najmniejszych kwadratów, uzyskane wówczas estymatory są zgodne i asymptotycznie nieobciążone. W literaturze przedstawione są także inne metody szacowa-

<sup>8</sup> Dokładniej własności funkcji autokorelacji oraz autokorelacji cząstkowej przedstawione są w pracach: Tsay (2002, s. 24-25), Lutkepohl, Kratzig (2004), Chatfield (2004).

<sup>9</sup> W praktyce rząd opóźnień  $p$  ustala się na poziomie nie przekraczającym 20% długości szeregu czasowego obejmującego  $n$  obserwacji.

nia parametrów, na przykład za pomocą układu równań Yule'a – Walkera. Metoda ta jest szczegółowo opisana w pracy Boxa, Jenkinsa z 1983 roku (patrz Box, Jenkins 1983, s. 63-65). Estymatory uzyskane za pomocą klasycznej metody najmniejszych kwadratów w pewnych sytuacjach mogą się różnić od tych uzyskanych za pomocą równań Yule'a – Walkera (Osińska 2006, Tsay 2002).

Na etapie weryfikacji sprawdza się własności reszt modelu, które powinny mieć własności białego szumu. Oznacza to tyle, że współczynniki autokorelacji reszt nie powinny różnić się istotnie od zera. W tym celu analizie poddaje się wykresy funkcji autokorelacji i autokorelacji cząstkowej dla szeregu reszt modelu. Wizualną analizę wykresów wspomnianych funkcji należy uzupełnić testem Ljunga – Boxa, w którym statystyka zbudowana jest według wzoru:

$$Q = (n - d) \sum_{k=1}^K r_e^2(k) \quad (7)$$

gdzie:

$r_e(k)$  – funkcja autokorelacji reszt,

$n$  – liczba wyrazów wyjściowego szeregu czasowego,

$d$  – rząd operatora różnic.

Zakładając, iż dobrany model jest odpowiedni statystyka ta ma rozkład  $\chi^2$  o  $K-p-q$  stopniach swobody.

Weryfikacja modelu powinna także obejmować badanie istotności parametrów modelu. Jeżeli model nie zostanie zweryfikowany pozytywnie należy wrócić do etapu pierwszego i ponownie przeprowadzić identyfikację. Pozytywnie zweryfikowany model wykorzystuje się do sporządzenia prognozy (Zeliaś i inni, 2004, Tsay 2002, Dudek 2005, Witkowska 2008).

Dokładność predykcji związana jest z oceną trafności prognozy określonej na podstawie błędów *ex post*. Celem tych mierników jest syntetyczny opis empirycznego rozkładu odchyłeń realizacji zmiennej prognozowanej od ustalonego poziomu prognozy, jaki osiągnięto w przeszłości na odcinku czasu, z którego pochodzą dane statystyczne. Warunkiem efektywnej budowy tych mierników jest posiadanie odpowiednio bogatych materiałów o skonstruowanych prognozach i odpowiadających im realizacji zmiennej prognozowanej. Mierniki te dostarczają informacji na temat tendencji występujących w błędach prognozy, a więc o ewentualnej dezaktualizacji modeli.

Pośród wielu prezentowanych w literaturze mierników jakości prognoz właściwym wydaje się zastosowanie następujących mierników: średni błąd

prognozy ( $ME - Mean Error$ )<sup>10</sup>, odchylenie standardowe błędów prognozy ( $RMSE - Root Mean Squared Error$ )<sup>11</sup>, średni bezwzględny procentowy błąd predykcji ( $MAPE - Mean Absolute Percentage Error$ )<sup>12</sup> oraz współczynnik Theila<sup>13</sup> ( $I^2$ ):

$$ME = \frac{1}{m} \sum_{t \in I_p} (y_t - y_t^*) \quad (8)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{t \in I_p} (y_t - y_t^*)^2} \quad (9)$$

$$MAPE = \frac{1}{m} \sum_{t \in I_p} \left| \frac{y_t - y_t^*}{y_t} \right| \cdot 100 \quad (10)$$

$$I^2 = \frac{mS_p^2}{\sum_{t \in I_p} y_t^2} \quad (11)$$

gdzie:  $m$  – liczba par obserwacji,

$y_t$  – rzeczywiste realizacje zmiennej prognozowanej,

$y_t^*$  – wartość prognozy dla zmiennej prognozowanej,

$I_p$  – okres empirycznej weryfikacji prognoz.

---

<sup>10</sup> Średni błąd prognozy służy do oceny, czy predykcja była obciążona, czy też nie. Jeśli mamy do czynienia z predykcją nieobciążoną, wartość średniego błędu prognozy oscyluje wokół zera. Duże odchylenia świadczą o tym, że zasada predykcji nieobciążonej nie została zachowana.

<sup>11</sup> Informuje o przeciętnym odchyleniu prognoz od wartości rzeczywistych w przedziale empirycznej weryfikacji prognoz.

<sup>12</sup> Miernik ten dostarcza informacji na temat, jaki procent rzeczywistej wartości zmiennej prognozowanej stanowi błąd bezwzględny, błąd ten nie jest symetryczny, wyżej ocenia przeszacowanie prognoz niż niedoszacowanie.

<sup>13</sup> Pierwiastek kwadratowy współczynnika Theila informuje jaki był przeciętny względny błąd prognozy bez względu na to, co było przyczyną takiego stanu rzeczy. W przypadku gdy współczynnik Theila przyjmuje wartość zero ( $I^2=0$ ) prognozy są idealnie trafne, im większa wartość tego współczynnika, tym większe są różnice między prognozami i wartościami rzeczywistymi.

Powyższe mierniki pozwalają rozstrzygnąć o dopuszczalności prognoz oraz o wielkości odchyień realizacji zmiennej prognozowanej od sformułowanych prognoz (Zeliaś 1997; Czerwiński, Guzik 1980).

## MODELE AUTOREGRESJI ORAZ PROGNOZY CEN PRODUKTÓW ROLNYCH

W opracowaniu zbadano kształtowanie się cen w skupie podstawowych produktów rolnych w Polsce, zbóż: pszenicy oraz żyta, żywca rzeźnego: trzody chlewnej oraz bydła<sup>14</sup>. Zakres czasowy badań obejmuje okres styczeń 1996 - maj 2010, co daje łącznie 173 obserwacje dla każdej analizowanej zmiennej, zostały zaczerpnięte z Biuletynów Statystycznych Głównego Urzędu Statystycznego.

Analizę szeregów cen rozpoczęto od zbadania ich stacjonarności za pomocą rozszerzonego testu Dickey'a - Fullera (z wyrazem wolnym oraz trendem)<sup>15</sup>. Wyniki testu dla poszczególnych zmiennych zamieszczono w tabeli 1. Na podstawie otrzymanych wartości statystyki ADF oraz odpowiadających im wartości prawdopodobieństwa *p-value*<sup>16</sup>, dla wszystkich rozważanych zmiennych odrzucono hipotezy zerowe o istnieniu pierwiastka jednostkowego. Tym samym przyjęto, że poszczególne zmienne są stacjonarne.

**Tabela 1. Wyniki rozszerzonego testu Dickey'a-Fuller'a dla przyjętych do badania zmiennych**

Zmienna	Wartość statystyki ADF	Wartość p-value
cena_pszenicy	-3,88	0,012
cena_żyta	-4,55	0,001
cena_trzody	-5,13	0,000
cena_bydła	-3,39	0,050

Źródło: obliczenia własne.

Po zbadaniu stacjonarności kolejnym krokiem w budowie modeli autoregresyjnych jest ustalenie rzędu opóźnień. Wykorzystuje się w tym celu

<sup>14</sup> Dane o cenach skupu dotyczą cen płaconych przez podmioty gospodarcze skupujące produkty rolne bezpośrednio od ich producentów. Przeciętne ceny obliczane są jako iloraz wartości (bez podatku od towarów i usług) i ilości poszczególnych produktów rolnych.

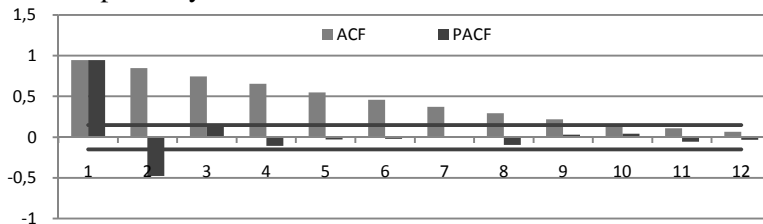
<sup>15</sup> Obliczenia wykonano w programie GRET.L.

<sup>16</sup> Przy weryfikacji hipotez za pomocą pakietów komputerowych ważne staje się wprowadzenie drugiego *ex post* poziomu istotności, oprócz zakładanego poziomu istotności  $\alpha$ . Ten drugi poziom istotności zwany jest „komputerowym poziomem istotności” lub „poziomem prawdopodobieństwa”, oznaczany jest przez *p-value*. Jeśli  $\alpha > p\text{-value}$  to na danym poziomie istotności  $\alpha$  odrzucamy hipotezę zerową.

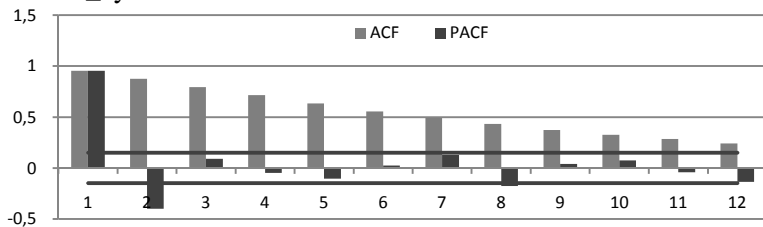
funkcję autokorelacji ACF oraz funkcję autokorelacji cząstkowej PACF. Analizę wykresów uzupełniono testem Quenouille'a. W teście tym wartości funkcji autokorelacji cząstkowej porównuje się z wartością krytyczną<sup>17</sup>. Na rysunku 1 przedstawiono wykresy funkcji autokorelacji ACF i autokorelacji cząstkowej PACF dla poszczególnych zmiennych, dodatkowo zaznaczono wartości krytyczne dla testu Quenouille'a.

**Rysunek 1. Wykres funkcji autokorelacji ACF i autokorelacji cząstkowej PACF dla poszczególnych zmiennych**

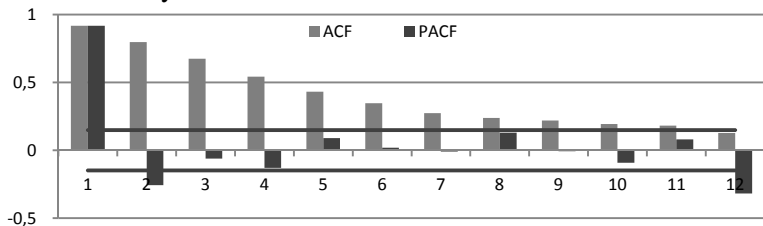
a) cena\_pszencicy



b) cena\_żyta

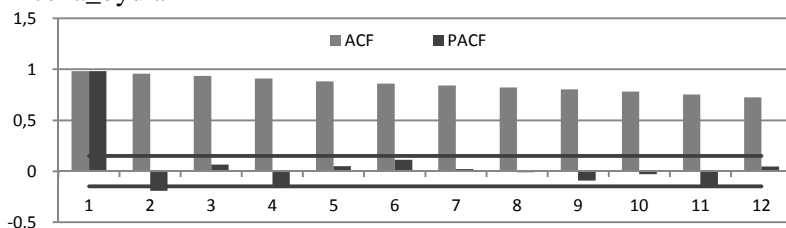


c) cena\_trzody



<sup>17</sup> Wartość krytyczna obliczana jest według formuły:  $\pm \frac{1,96}{\sqrt{T}}$ , gdzie T jest liczbą obserwacji.

d) cena\_bydła



Źródło: Opracowanie własne.

Na podstawie wartości funkcji autokorelacji cząstkowej oraz wartości krytycznej statystyki testu Quenouille’a dla poszczególnych zmiennych ustalono następujące rzędy opóźnień: cena\_pszonicy – 2, cena\_żyta – 8, cena\_trzody – 12, cena\_bydła – 4. Uwzględniając jednocześnie stopień zintegrowania oraz rząd opóźnień dla analizowanych zmiennych, dalszej analizie będą poddawane następujące modele: cena\_pszonicy – AR(2); cena\_żyta – AR(8); cena\_trzody – AR(12); cena\_bydła – AR(4).

Parametry modeli zostały oszacowane za pomocą klasycznej metody najmniejszych kwadratów. Równania modeli wraz z błędami szacunku parametrów przedstawione są poniżej. Zapisane równania zawierają tylko te parametry, które okazały się istotne:

$$\text{Cena\_pszonicy} \quad \hat{y}_t = 4,00 + 1,41 y_{t-1} - 0,49 y_{t-2}$$

(1,16) (0,07) (0,07)

$$\text{Cena\_żyta} \quad \hat{y}_t = 2,48 + 1,44 y_{t-1} - 0,63 y_{t-2} + 0,35 y_{t-7}$$

(0,93) (0,08) (0,14) (0,14)

$$\text{Cena\_trzody} \quad \hat{y}_t = 0,38 + 1,19 y_{t-1} - 0,27 y_{t-2} - 0,21 y_{t-7}$$

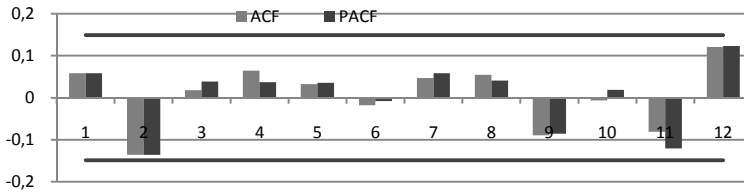
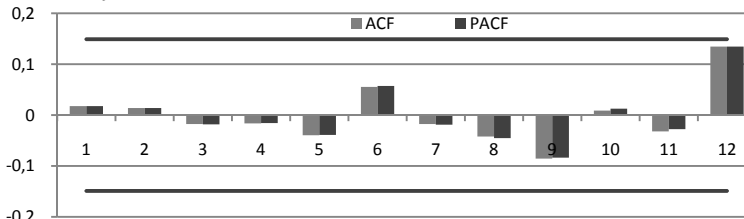
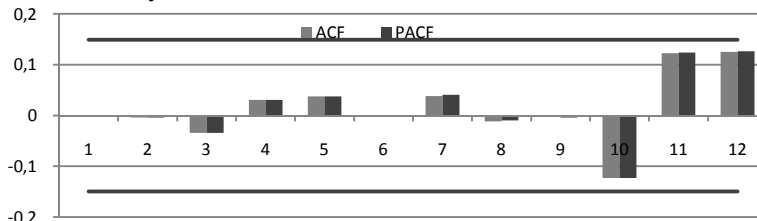
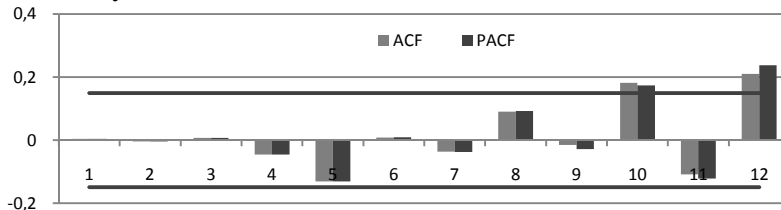
(0,14) (0,08) (0,12) (0,12)

$$\text{Cena\_bydła} \quad \hat{y}_t = 1,28 y_{t-1} - 0,44 y_{t-2} + 0,22 y_{t-3}$$

(0,07) (0,12) (0,12)

Kolejnym etapem weryfikacji modeli było zbadanie własności reszt modeli, w tym celu przeanalizowano wartości funkcji autokorelacji oraz funkcji autokorelacji cząstkowej reszt modeli (rysunek 2)<sup>18</sup>.

<sup>18</sup> W praktyce na wykresach autokorelacji i autokorelacji cząstkowej zaznacza się granice wyznaczone przez podwojone wartości błędów standardowych odpowiednio funkcji ACF i PACF. Za wartości różne od zera uznawane są te współczynniki, które przekraczają zaznaczone granice (Dudek 2005).

**Rysunek 2. Wykres funkcji autokorelacji ACF i autokorelacji cząstkowej PACF dla reszt modeli dla poszczególnych zmiennych**a) *cena\_pszenyicy*b) *cena\_żyta*c) *cena\_trzody*d) *cena\_bydła*

Źródło: obliczenia własne.

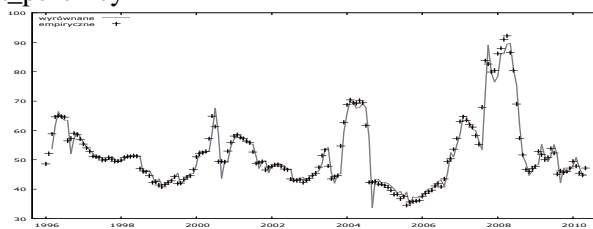
Ostatecznie modele przeszły pozytywnie etap weryfikacji. Mogą być zatem wykorzystane do celów prognostycznych. Na rysunku 3 zaprezentowano wartości rzeczywiste oraz wartości wyrównane wyznaczone na podstawie oszacowanych modeli. Wykresy wskazują na dobre dopasowanie modeli do danych empirycznych, potwierdzają to również wartości współczynników determinacji, które wynoszą odpowiednio: 0,92 dla modelu dla



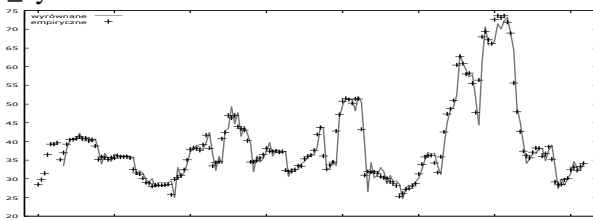
zmiennej cena\_pszenicy, 0,93 dla zmiennej cena\_żyta, 0,87 dla zmiennej cena\_trzody oraz 0,98 dla modelu dla zmiennej cena\_bydła.

**Rysunek 3. Empiryczne i wyrównane za pomocą modeli wartości poszczególnych zmiennych w okresie styczeń 1996 – maj 2010**

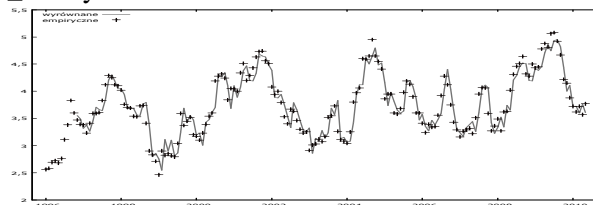
a) cena\_pszenicy



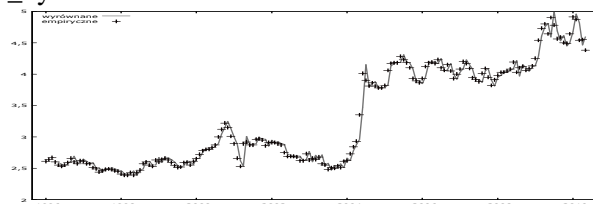
b) cena\_żyta



c) cena\_trzody



d) cena\_bydła



Źródło: obliczenia własne.

Po oszacowaniu modeli oraz wstępnym sprawdzeniu ich jakości dokonano oceny dokładności prognoz. W tym celu obliczono mierniki *ex post* dokładności predykcji: średni błąd prognozy (*ME*), odchylenie standardowe błędów prognozy (*RMSE*), średni bezwzględny procentowy błąd predykcji

(MAPE) oraz współczynnik Theila ( $I^2$ ). Wartości miar dla poszczególnych modeli przedstawia tabela 2.

**Tabela 2. Miary dokładności prognoz dla oszacowanych modeli dla poszczególnych zmiennych**

Zmienna	ME	RMSE	MAPE	$I^2$	$I$
cena_pszenicy	0,00	3,21	3,87%	0,83	0,91
cena_żyta	0,00	2,70	4,67%	0,85	0,92
cena_trzody	0,00	0,19	4,29%	0,83	0,91
cena_bydła	0,00	0,10	2,13%	0,90	0,95

Źródło: Obliczenia własne.

Porównując wartości poszczególnych miar dla kolejnych produktów wnioskujemy, że uzyskane prognozy mogą być uznane za bardzo dokładne, świadczy o tym fakt, że średni błąd prognozy w całym analizowanym okresie wynosi zero. Prognozowane wartości cen różniły się od rzeczywistych średnio o 3,21 PLN/100 kg w przypadku pszenicy; 2,70 PLN/100 kg – żyta, 0,19 PLN/1kg – trzody i 0,10 PLN/1kg w przypadku bydła. Wartości te mogą wydawać się duże w szczególności dla zmiennych cena\_pszenicy i cena\_żyta, ale porównując je z średnimi cenami w skupie (wynoszą one odpowiednio: pszenica – 51,94 PLN/100 kg, żyto – 39,16 PLN/100 kg, trzoda – 3,73 PLN/100 kg, bydło – 3,29 PLN/100 kg) stanowią one niewielki ułamek tej ceny. Kolejna miara dokładności prognoz – średni bezwzględny procentowy błąd predykcji (która informuje o procentowych odchyleniach prognozy od rzeczywistej wartości zmiennej prognozowanej dla kolejnych okresów czasu) wskazuje, że prognozowane ceny w okresie weryfikacji różnią się od wartości rzeczywistych o mniej niż 5%. Tak niska wartość MAPE pozwala na uznanie prognoz za dopuszczalne.

**Tabela 3. Prognozy cen produktów rolnych na 2010 rok**

Wyszczególnienie	VI	VII	VIII	IX	X	XI	XII
cena pszenicy	48,72	49,69	50,32	50,74	51,02	51,22	51,35
cena żyta	34,97	35,21	35,82	36,77	37,14	37,45	37,77
cena trzody	3,82	3,82	3,86	3,86	3,81	3,83	3,79
cena bydła	4,3	4,3	4,29	4,28	4,28	4,28	4,27

Źródło: Obliczenia własne.

Za pomocą oszacowanych modeli wyznaczono prognozy cen analizowanych produktów rolnych w skupie na okres czerwiec 2010 – grudzień 2010 rok, patrz tabela 3.

## ZAKOŃCZENIE

Prowadzenie gospodarstwa rolnego w warunkach rynkowych, związane jest z ciągłym podejmowaniem decyzji różnego rodzaju, które mają pośredni lub bezpośredni wpływ na dochody z prowadzonej działalności. Dynamicznie zmieniające się otoczenie gospodarstw rolnych oraz rosnąca konkurencja powodują, że wybory dokonywane są z niepewnością a skutki tych wyborów rozciągają się na wiele okresów. Konieczność przygotowania się na przyszłe skutki dzisiejszych działań wymusza na osobach podejmujących decyzje konieczność sporządzania prognoz gospodarczych, a na ich podstawie planów działania. Właściciel gospodarstwa rolnego, zajmujący się czy to hodowlą zwierząt, czy też uprawą zbóż, chce znać przyszłe ceny produktów rolnych aby móc odpowiednio wcześniej podjąć odpowiednie decyzje produkcyjne. Zatem głównym celem prognozowania w takim przypadku jest wskazanie najbardziej prawdopodobnego przebiegu w przyszłości zjawisk i procesów gospodarczych oraz ewentualnych skutków podjętych działań. W pracy zaprezentowano możliwość wykorzystania modeli autoregresyjnych do prognozowania cen wybranych produktów rolnych w skupie. Z analizy wynika, że modele te są efektywnym narzędziem krótkookresowego prognozowania cen. Efektywność oszacowanych modeli przekłada się na dobre dopasowanie modeli do danych empirycznych – świadczą o tym wysokie wartości współczynników determinacji (powyżej 0,85) oraz odpowiednio niskimi wartościami błędów *ex post*. Wyznaczone prognozy na okres czerwiec 2010 – grudzień 2010 utrzymują się w ogólnej tendencji, ceny zbóż nieznacznie wzrosną natomiast ceny żywca nie ulegną większym zmianom.

## LITERATURA

- Biuletyn Statystyczny z lat 1996-2010, Główny Urząd Statystyczny, Warszawa
- Box G.E. P., Jenkins G. M. (1983), *Analiza szeregów czasowych. Prognozowanie i sterowanie*, Państwowe Wydawnictwo Naukowe, Warszawa.
- Charemza W. W., Deadman D. F. (1997), *Nowa ekonometria*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Chatfield Ch. (2004), *The analysis of Time Series. An Introduction*, CRC Press LLC, Florida USA.
- Czerwiński Z., Guzik B. (1980), *Prognozowanie ekonometryczne. Podstawy teoretyczne i metody*, Państwowe Wydawnictwo Naukowe, Warszawa.

- Dudek H. (2005), *Prognozowanie cen skupu mięsa drobiowego za pomocą sezonowego modelu ARIMA*, „Roczniki Naukowe Stowarzyszenia Rolnictwa i Agrobiznesu”, tom VII, zeszyt 5.
- Gruszczyński M., Podgórska M. (2004), *Ekonometria*, Szkoła Główna Handlowa w Warszawie, Warszawa.
- Hamilton J. D. (1994), *Time Series Analysis*, Princeton University Press, New Jersey USA.
- Hatanaka M. (2004), *Time Series Based Econometrics. Unit Roots and Cointegration*, Oxford University Press, New York USA.
- Kwaśnicki W. (2010), *Cykl świński – próba modelowania i analizy*, <http://prawo.uni.wroc.pl/~kwasnicki/todownload/Cykl%20swinski.pdf>, z dn. 20.07.2010.
- Lutkepol H., Kratzig M. (2004), *Applied Time Series Econometrics*, Cambridge University Press, New York USA.
- Majewski J. (2006), *Ceny skupu mleka w Polsce – Analiza i prognozowanie*, „Roczniki Naukowe Stowarzyszenia Rolnictwa i Agrobiznesu”, t. VIII, zeszyt 2.
- Osińska M. (2007), *Współczesna ekonometria*, Dom Organizatora, Toruń.
- Osińska M. (2006), *Ekonometria finansowa*, Polskie Wydawnictwo Naukowe, Warszawa.
- Sobczyk M. (2008), *Prognozowanie. Teoria, przykłady, zadania*, Wydawnictwo Placet, Warszawa.
- Stępień S. (2009), *Klasyczne i współczesne teorie cyklu świńskiego – wnioski dla gospodarstw trzodowych*, „Roczniki Naukowe Stowarzyszenia Rolnictwa i Agrobiznesu”, tom XI, zeszyt 3.
- Tłuczak A. (2008a), *Metody prognozowania cen na rynku mięsa*, [w:] S. Sokołowska, A. Bisaga (red.), *Wieś i rolnictwo w okresie zmian, Problemy funkcjonowania i rozwoju rolnictwa*, Wydawnictwo Uniwersytetu Opolskiego, Opole.
- Tłuczak A. (2008), *Prognozowanie cen mięsa wieprzowego z zastosowaniem modelu Wintersa*, „Prace Komisji Naukowych”, Zeszyt nr 32, PAN, Oddział w Katowicach, Katowice.
- Tłuczak A., *Efektywność modeli adaptacyjnych w prognozowaniu cen rolnych*, [w:] R. Żelazny (red.), *Koniunktura gospodarcza a funkcjonowanie rynków*, Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice.
- Tsay R. S. (2002), *Analysis of Financial Time Series. Financial Econometrics*, John Wiley & Sons INC., New York, USA.
- Witkowska D., Matuszewska A., Kompa K. (2008), *Wprowadzenie do ekonometrii dynamicznej i finansowej*, wydawnictwo SGGW, Warszawa.

- Zeliaś A., Pawełek B. (2004), Wanat S., *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa.
- Zeliaś A. (1997), *Teoria prognozy*, Polskie Wydawnictwo Ekonomiczne, Warszawa.

