Development of Milk Prices in the Czech Republic

Ondřej ŠIMPACH[1], Marie ŠIMPACHOVÁ PECHROVÁ[2]

[1] University of Economics, Prague, Czech Republic
ondrej.simpach@vse.cz
[2] Institute of Agricultural Economics and Information, Prague, Czech Republic
simpachova.marie@uzei.cz

Abstract. The sector of milk production underwent through the crisis. Knowing the future price can help the management of the farm to minimize the risks related to the decision-making in production. Hence, the aim of the paper is to find optimal model for modelling of monthly prices of milk from 01/1998 to 06/2016 and predict them for 07/2016–06/2017.
First, the development of milk producers' prices is modelled by Autoregressive Integrated Moving Averages models. Consequently, the relations with monthly prices of maize (02/2006–06/2016) were examined using Vector Autoregressive (VAR) model. ARIMA model suggests prices of milk in the interval from 5.97 CZK/l to 7.06 CZK/l that is more realistic prediction than in case of VAR model that predicts lower prices (5.86–6.38 CZK/l). There might be missing some important variables in VAR model of prices of milk that can improve the prediction capability. Therefore, the challenge for future research is to improve the models further using regression analysis with relevant determinants.

Keywords: Milk Price, ARIMA Model, Vector Autoregressive Model

1 Introduction

The sector of milk production underwent a crisis recently. The price has always been volatile, but "Extreme price spike and volatility in agricultural commodity prices creates negative effects on macroeconomic instability, posing a threat to food security in many countries," [4]. Like in other sectors, the price is determined based on the agreement between the seller and the buyer. However, as Jaile-Benitez, Ferrer-Comalat and Linares-Mustarós [5] noted, "this agreement rarely remains stable because it is often reached after unwanted pressures, creating situations of dissatisfaction that involve one of the two parties". The market situation is unfavourable for the farmers as they are not unified and their bargaining power is lower than the power of dairy companies. The reasons for asymmetric price transmission in the agro-food chain was examined for example by Bakucs, Falkowski and Fertö [3] using meta-analysis of existing studies.

In the US, there might be a problem with speculative bubbles at stock markets. Adämmer and Bohl [1] found that they were present in wheat prices in years 2003–2013, but for prices of corn and soybeans the results were inconclusive.

Besides, the competition with other producers on the market is an important reason for prices fluctuations. Especially after the entrance to the EU, impacts of the single market on the Czech milk sector were significant as about one third of Czech milk production is being exported. Cancelling the import tariffs lead to increased competition on the market. The producer's prices were low and the costs of production high [11]. The important changes since the entrance to the EU happened also in the size of the farms – especially those aimed at pig production. The number of farms has reduced, but their size has increased. This enabled them to adjust their production and achieve the returns to scale. [7].

Modelling and prediction of prices and finding the relations between the developments of various prices had been a subject of examination by many authors. Rumánková [8] used Box-Jenkins Autoregressive Integrated Moving Averages (ARIMA) modelling method to project the prices of selected agricultural commodities. She found that mostly the time series are integrated by the order of 1. Saengwong et al. [9] also found that the prices of broilers, cattle, duck and hogs are stationary when their first differences were taken into account. Šimpach [10] modelled honey prices.

Regarding the prediction capability of the various models, Tenege and Kuchler (1994) compared and evaluate two variations of the present-value model (PV1 and PV2), an ARIMA, a vector autoregression (VAR), and an error-correcting model (ECM) based on root mean-squared error (forecast accuracy) and Henriksson-Merton test (ability to predict turning points).

In the article, in order to examine the current development and to project the future development of the milk price in the Czech Republic our paper uses both, univariate and multivariate approaches to the time series analysis. Particularly, it scrutinises and forecasts individual time series by Box-Jenkinson methodology (ARIMA models), and examines the type of the relations among multiple time series using VAR model. Both models are also used for predictions.

2 Methodology

There are basically two groups of methods used for the time series analysis. In the study, we employ both, the univariate and multivariate. First, the Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) analysis is applied on time series of individual series of price of milk. Second, the influence of the prices of the feed maize on the commodity prices (short-term relationship) was examined using Vector Autoregressive (VAR) model. Both models are diagnostically tested. At the beginning, the time series are examined by Augmented Dickey-Fuller (ADF) test whether they are stationary or non-stationary, because VAR model can be used only if they are integrated of the same order. We applied ADF test without constant and trend.

$$\Delta Y_t = \beta\, Y_{t-1} + \sum_{i=1}^{m} \alpha_i Y_{t-i} + \varepsilon_t \qquad (1)$$

where ΔY_t is the first difference of the examined variable, t is time, m is the maximum length of the lagged dependent variable, α, β are parameters, and ε_t is a pure white noise error term.

Diagnostic of the type of ARIMA model is done by Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) that are plotted to determine the order p of Autoregressive (AR) process and order q of Moving Average (MA) process. Sample ACF is calculated from time series Y_t.

$$ACF(k) = \frac{\sum_{t=k+1}^{n}(Y_t - \overline{Y})(Y_{t-k} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} \qquad (2)$$

ACF and PACF can be expressed graphically in correlograms. PACF consider the fact that correlation between two random variables is often caused by the fact that those two variables are correlated with the third variable and is adjusted from the influence of other quantities. PACF is defined for k >2 as

$$\begin{aligned} PACF(1) &= cor(Y_{t+1}, Y_t) \\ PACF(k) &= cor(Y_{t+k} - P_{t,k}(Y_{t+k}), Y_t - P_{t,k}(Y_t)) \end{aligned} \qquad (3)$$

where $P_{t,k}(Y)$ denotes the projection of Y onto space spanned by Y_{t+1}, ..., Y_{t+k-1}.

Consequently, the appropriate type of the model is identified. There are compared the results of Autoregressive (AR) model, Moving Average (MA) model, Autoregressive Integrated Moving Average (ARIMA) model in terms of statistically significant parameters and their ability to explain the correlation structure of the process that generated the time series. Then the ARIMA(p, d, q) model can be written in the form

$$Y_t = \beta + \sum_{i=1}^{p}\alpha_i Y_{t-i} + \sum_{j=1}^{q}\delta_j \varepsilon_{t-j} \qquad (4)$$

After finding the appropriate model, the presence of autocorrelation is checked using Breusch-Godfrey serial autocorrelation LM test. Null hypothesis states that there is no serial autocorrelation. If the calculated value of the test exceeds the tabled test criterion from Fisher and χ^2 distribution the null hypothesis is rejected and there is autocorrelation. Durbin-Watson test also enables to decide whether there is autocorrelation. However, it has some drawbacks (e.g. it has “grey” zones where it is not possible to evaluate the test). Its value should be between lower and upper level of the critical values of Durbin-Watson distribution. Values around 2 indicate that there is no autocorrelation.

Heteroscedastity is tested by Autoregressive Conditional Heteroscedasticity (ARCH) test. Null hypothesis is again that there is no heteroscedasticity present (i.e. that the variance of the residues is constant and finite). The test is also using Fisher and χ^2 critical values and rejects the null hypothesis if the calculated value of the test exceeds the table values.

Normality is tested by Jarque-Bera test with null hypothesis that the residues are normally distributed. Calculated value of the test is compared to critical value of

Jarque-Bera distribution. If the value exceeds the table one, null hypothesis is rejected. All tests are done at 0.05% level of significance.

Consequently, the fitted models are used to predict the future producers' prices of milk in future 12-month period. Also 95% confidence intervals are elaborated.

Second, the possible correlation between the time series of milk and feed maize is examined. It is supposed that variables have similar trend. They also have to be integrated of the same order. Therefore, they are tested by ADF test (after and before seasonal adjustment). Consequently, the time series are tested for spurious regression by Granger test. General VAR(p) model can be written in the form

$$\mathbf{Y}_t = \boldsymbol{\beta} + \sum_{i=1}^{m} \boldsymbol{\Phi}_i \mathbf{Y}_{t-i} + \boldsymbol{\varepsilon}_t \qquad (5)$$

where $\boldsymbol{\beta}$ is l x 1 dimensional vector of constants, $\boldsymbol{\Phi}_i$, i = 1, 2, …, m are l x l dimensional non-random matrices of AR parameters and ε_t is l-dimensional process of white noise. In our case, price of milk is modelled by VAR with two time series (in natural logarithms).

$$\ln Y_{1,t} = c_1 + \ln Y_{1,t-1} + \ln Y_{1,t-2} + \ln Y_{1,t-3} + \ln Y_{2,t-1} + \ln Y_{2,t-2} + \ln Y_{2,t-3} + \varepsilon_{1,t}$$
$$\ln Y_{2,t} = c_2 + \ln Y_{1,t-1} + \ln Y_{1,t-2} + \ln Y_{1,t-3} + \ln Y_{2,t-1} + \ln Y_{2,t-2} + \ln Y_{2,t-3} + \varepsilon_{2,t} \qquad (6)$$

Then the model was econometrically verified. Autocorrelation was tested by residual serial correlation LM tests and heteroscedasticity by VAR residual heteroskedasticity tests. Using Jarque-Bera test it was examined whether the distribution of residuals is normal. The projections with 95% confidence intervals are done for 12-month period.

3 Data

Data were taken from Agris [2]. Originally, the frequency of milk price was weekly, but was transformed on monthly data by calculating the monthly average as the prices of feed were available on the monthly bases. Observed period is from 01/1998 to the latest available data for all commodities – 06/2016. There were 222 observations. Average annual prices are given in Table 1. Calculations are done in EViews 8.

Table 1. Average annual producers' prices of milk, and fodder maize. Source: own elaboration.

Year	**1998**	**1999**	**2000**	**2001**	**2002**
Milk [CZK/l]	7.88	7.2	7.45	7.73	8.09
Year	**2003**	**2004**	**2005**	**2006**	**2007**
Milk [CZK/l]	7.8	7.94	8.25	7.86	8.22
Fodder maize [CZK/t]				2921.09*	4182.92
Year	**2008**	**2009**	**2010**	**2011**	**2012**
Milk [CZK/l]	8.76	6.2	7.39	8.28	7.79
Fodder maize [CZK/t]	4634.17	2800.00	3282.25	4797.64	4640.54

Year	2013	2014	2015	2016**
Milk [CZK/l]	8.4	9.51	7.86	6.93
Fodder maize [CZK/t]	4975.33	4252.67	3781	4085.57

4 Results

First, the development of the time series was scrutinized. Second, the development was modelled individually by Box-Jenkinson methodology (by ARIMA model). Finally, a VAR model was constructed and utilized for price predictions.

The time series was plotted to get the basic idea about the character of the development of prices (see Fig. 1). The highest increase in milk price was in 01/2008 (10.10 CZK/l). It was increasing since 11/2007, but after reaching the maximum at the beginning of the next year, it decreased again below 9 CZK/l in 06/2018. The reason for increase was the lack of milk on the EU market. However, the lack was replaced by the surplus later and the prices decreased again. “Clear explanation, why there is suddenly a lot of milk while there was lack of it last year, does not exists. The only sure thing is that the production of important exporters such as Australia or New Zealand has started again and it began to remain more milk in Europe. [6] Nevertheless, the year with highest average price of milk was 2014, when the monthly prices were moving over 9 CZK/l. On the other hand, minimal prices were noted in 2009 (6.20 on average) when they decreased even on 5.92 CZK/l in 08 and 09/2009. In over 50% of cases, the prices were higher than 7.88 CZK/l, 25% months were the prices lower than 7.50 CZK/l and higher than 8.21 CZK/l.

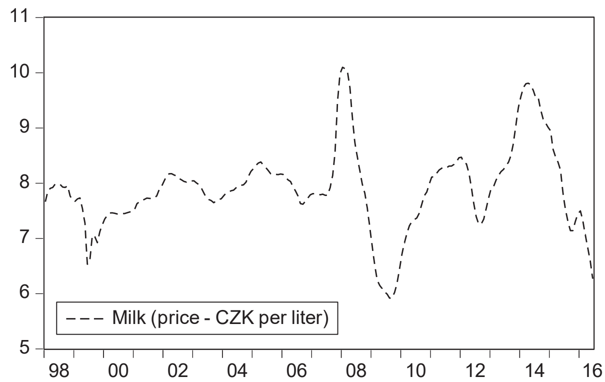

Fig. 1. Development of prices of milk (01/1998-06/2016). Source: own elaboration.

The price of milk was modelled individually by Box-Jenkins methodology and was further used for prediction. The results of the model together with verification tests are displayed in Table 2. Fisher seasonality test (F-test) revealed that time series is seasonal, but seasonal parameter (SAR or SMA) is not statistically significant. Therefore, it is not modelled. After testing, it was found that the most suitable model for price of milk was ARIMA(2,0,1)c. It means that the time series was stationary at the 5% significance level (the degree of integration is 0) and that the prices depends on the price one and two months ago (AR process is of degree 2) and that the MA model is of 1st order.

There was added a dummy variable IMP that takes value of 1 in May 1999 and zero in other case in order to capture the shock in the time series. Otherwise the residues would be autocorrelated. Constant 7.78 equals almost to mean dependent variable (7.90). Durbin-Watson test suggests that there is not autocorrelation (its value is almost equal to 2). Breusch-Godfrey test confirms that there is no autocorrelation at 5% level of significance (the probability of Fisher and χ^2 is higher than 0.05). The distribution of residues is not normal, finite and constant as the time series is long (18 years) and variable. Due to the presence of heteroscedasticity, we used Heteroscedasticity and Autocorrelation Corrected (HAC) errors. These standard errors take into account the heteroscedasticity and results are not biased (undervalued).

Table 2. Price of milk: ARIMA(2,0,1)c model, impulse = 1 (05/1999). Source: own elaboration.

Variable	**Coeff. (Std. error) Prob.**	**Model diagnostic**	
β	7.7834 (0.2539) ***	R^2	0.9842
IMP	0.2652 (0.0421) ***	Adjusted R^2	0.9840
AR(1)	1.6147 (0.1195) ***	F-statistic	3357.5600
AR(2)	-0.6508 (0.1152) ***	Prob. (F-statistic)	0.0000
MA(1)	0.3883 (0.1232) ***	Durbin-Watson statistic	2.0217
Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	2.1725		
Prob. $F_{[2,215]}$	0.1164		
Obs* R^2	4.3980		
Prob. $\chi^2_{[2]}$	0.1109		

Note: *** marks statistical significance at α = 0.01

Fig. 2 presents the development of real prices of milk (drawn with dotted line) and prices fitted by the model (displayed with dashed line). The residues (difference between real and fitted values) are displayed by solid line.

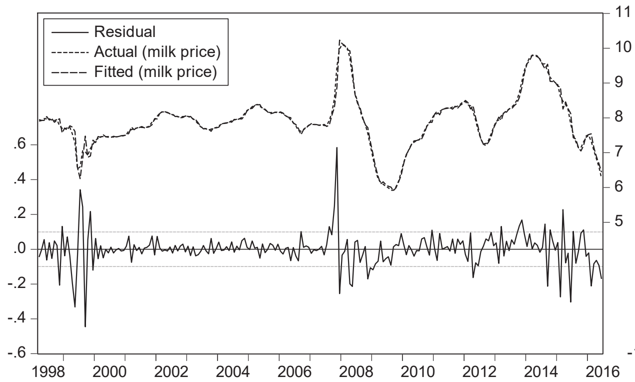

Fig. 2. Development of milk prices with residuals (01/1998-06/2016). Source: own elaboration.

Estimated model was used for predictions and construction of 95% confidence intervals. The results are displayed at Fig. 3. Despite that the prices of milk were decreasing at the end of the observed period (since 01/2016), the model suggests that there will be an increase again. At the beginning, the price shall be the lowest (5.97 CZK/l in 08/2017 and 5.98 CZK/l in 09/2017) and then increase up to 7.06 CZK/l in 06/2017.

Lower bound of 95% confidence interval predict further decrease of the price, down to 4.99 CZK/l in 12/2016, but then also the price shall be higher (5.27 CZK/l in 06/2017). Upper bound suggests that the price might increase up to 8.85 CZK/l in 06/2017.

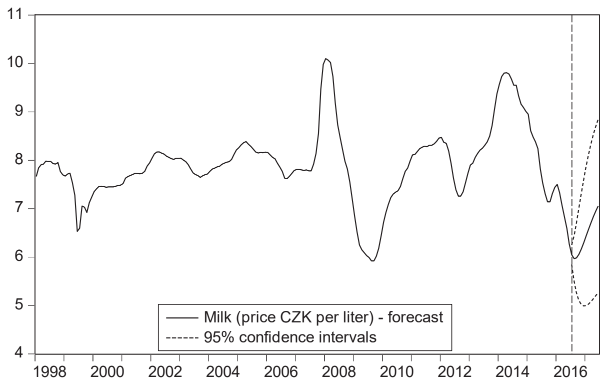

Fig. 3. Predictions of producers' prices of milk based on ARIMA model. Source: own elaboration.

VAR models examine the short-time relationship between the development of producers' prices of milk and possible determinants. In order to make possible the construction of the VAR models, the time series has to be integrated of the same order. While individual time series of the price of milk is modelled as non-seasonal, it enters the VAR model as seasonally adjusted. Also price of maize is seasonal and is seasonally adjusted. Milk price and fodder maize were tested by ADF test without constant and trend. It was found out that they are non-stationary and integrated of the order 1 (stationary after first differences). Granger spurious regression test revealed that the relation is regression (it is not spurious, hence, it is possible to further model the time series in VAR). It was found that fodder maize Granger cause price of milk. On the other hand, the null hypothesis that price of milk does not Granger cause the price of fodder maize was not rejected. Therefore, the relation and causality between the prices is according to the expectations. Then the time series are used in two-equation model and short-term relation is searched.

Results of the VAR model are presented in Table 3. Time series was included in logarithms (in order to improve the diagnostic tests). According to information criteria and FPE (Final Prediction Error) it is optimal to choose VAR model of second order. However, the diagnostics tests of this type of model were not optimal.

Therefore, we elaborated VAR model with three lags of milk and fodder maize (VAR model of the third order). Milk is statistically significantly dependent on itself in the first and second lag. It is also influenced by the prices of fodder maize one and two months ago. There were six unit impulses (dummy variable taking the value of 1 in case of the impulse, 0 otherwise) in periods 11/2007, 08, 10, 12/2012, and 10, 11/2015 to capture the shocks in the price development and to improve the econometric characteristics of the model. Price of the fodder maize is not dependent at the price of the milk. It depends only on its value one month ago and dummy variable IMP. While in the case of price of milk the development is statistically significantly explained by the constant at 0.1 level of significance, in the case of price of fodder it is already at 0.01 level of significance. VAR Residual serial correlation LM tests show that there is no serial correlation of the residues. Similarly, there was no heteroscedasticity present. However, due to the fact that time series is long, the distribution of the residues is not normal. This does not affect the mean prediction, but can have consequence in the prediction of confidence intervals (they can be biased).

Table 3. Estimates of VAR model of price of milk and fodder maize. Source: own elaboration.

	Price of milk (log)	**Price of maize (log)**
Variable	**Coeff. (Std. error)** $^{Prob.}$	**Coeff. (Std. error)** $^{Prob.}$
β	0.0548 (0.0359) *	1.0132 (0.2689) ***
Price of milk (log) (-1)	1.7710 (0.0917)***	0.5672 (0.6864)
Price of milk (log) (-2)	-0.7130 (0.1727)***	-0.1479 (1.2919)
Price of milk (log) (-3)	-0.0786 (0.0897)	-0.3196 (0.6713)
Price of maize (log) (-1)	0.0293 (0.0118) ***	0.8327 (0.0882) ***
Price of maize (log) (-2)	-0.0337 (0.0156) **	0.0529 (0.1167)
Price of maize (log) (-3)	0.0026 (0.0119)	-0.0333 (0.0891)
IMP (dummy)	0.0147 (0.0045) ***	0.1608 (0.0339) ***
Model diagnostic		
R^2	0.9936	0.8940
Adjusted R^2	0.9932	0.8875

F-statistic	2538.521	137.3163
Akaike info criterion	-6.2419	-2.2168
Schwarz criterion	-6.0580	-2.0329

Note: *** marks significance at α = 0.01, ** at α = 0.05, * at α = 0.1

VAR model enables to predict both prices – of fodder maize and milk. The development of the original data with joined projection can be found in Figure 4. Price of milk will further decrease on 5.85 CZK/l in 11/2016 when it will be the lowest. Than it will increase again and in 04/2017 and 07/2016 will be the same (6.13 CZK/l 6.14 CZK/l, respectively). In 06/2017 it shall achieve 6.38 CZK/l that is higher than it was in 06/2016. Upper bound of the confidence interval follows the increase of the price of milk, but without mild decrease as in the mean prediction. The price shall continue to increase since 07, 08/2016 when it is 6.27 CZK/l up to 8 CZK/l at the end of the projection horizon 06/2017. Lower bound suggests steep decrease down to unreal value of 4.74 CZK/l (05/2017).

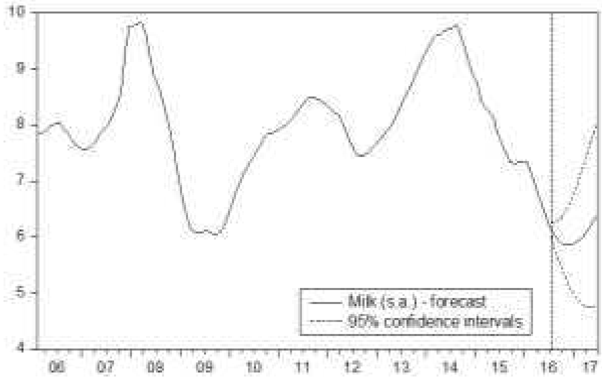

Fig. 4. Predictions of producers' prices of milk based on VAR model. Source: own elaboration.

Comparison of univariate and multivariate approaches can be seen from the Fig. 5. ARIMA and VAR models give almost similar results of the future development of prices of milk – i.e. the price will be lower at first (5.97 CZK/l in 08/2016 in the case of ARIMA and 5.85 CZK/l in 11/2016 in the case of VAR) and then it will increase up to 7.06 CZK/l, 6.38 CZK/l, resp., in 06/2017.

According to ARIMA, lower bound of 95% confidence interval predicts decrease of the price (down to 4.99 CZK/l in 12/2016) followed by increase up to 5.27 CZK/l in 06/2017. VAR's lower bound suggests only decrease of the price that should be as low as 4.76 CZK/l. This price is too low and in the context of current development of the price of milk is not real. In other words, regarding the pessimistic variant of price of milk development, ARIMA model gives better results, but also not realistic as the minimal price for the whole period was only 5.92 CZK/l. Upper bound of ARIMA model says that price might increase up to 8.85 CZK/l in 06/2017. VAR model also project constant increase of price, but only up to 8.00 CZK/l. This development is realistic, as during the examined period, the price of milk increased up to 10.10 CZK/l.

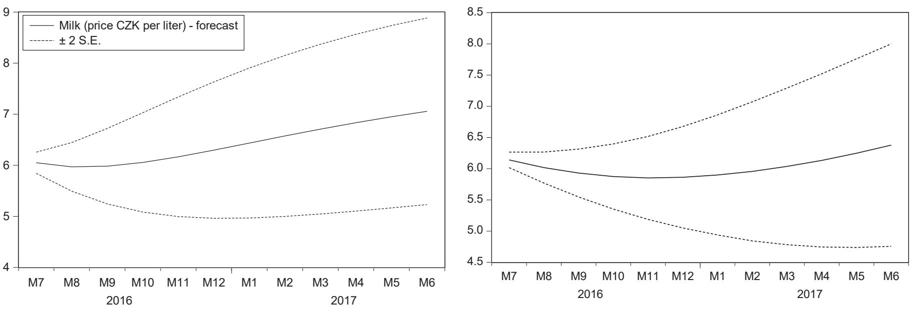

Fig. 5. Predicted prices of milk by Box-Jenkins methodology (left) and VAR model (right). Source: own elaboration.

Lower bound of ARIMA prediction of price of milk was much lower than VAR until the end of 2016. Then the VAR model predicts unrealistic low values. While in the case of prices of the milk, the difference between lower and upper bound starts at 0.41 CZK/l and finishes at 4.58 CZK/l – that is 2.44 CZK/l on average, in case of VAR it is only 1.77 CZK/l average difference between the highest and lowest price.

5 Conclusion

Analysis of the characteristics of agricultural product price volatility and trend forecasting are necessary to formulate and implement business strategies of agricultural holdings and for policy-making. In 2016, the sector of milk production underwent through the crisis. Therefore, the aim of the paper was to find the optimal model for modelling and predictions of monthly price of milk. It was modelled by Autoregressive Integrated Moving Averages (ARIMA) models (time series from 01/1998 to 06/2016) and by Vector Autoregressive (VAR) model in relation with price of maize (data from 02/2006 to 06/2016). Predictions are done for 12 months (until 06/2017).

VAR model predicts lower prices of milk in the pessimistic variant than could be in reality. Mean development and prices in upper bound of confidence interval are also lower, but feasible. The price of milk shall increase 7.06 CZK/l (ARIMA), 6.38 CZK/l (VAR) at the end of predicted period, but will experience the decrease under 6.00 CZK/l at the end of 2016. There might be missing some important variables in VAR model of prices of milk that can improve its prediction capability. Therefore, the challenge for future research is to find other time series that can be included in the VAR model of price of milk.

Acknowledgements. The research was supported by the Czech Science Foundation project no. P402/12/G097 DYME – "Dynamic Models in Economics" and from Internal Research Project no. 1277 of Institute of Agricultural Economics and Information.

References

1. Adämmer P., Bohl M. T.: Speculative bubbles in agricultural prices. The Quarterly Review of Economics and Finance, 55, 67–76 (2015).
2. Agris: Agrarian www portal. Available at http://www.agris.cz/ last accessed Nov 1, 2016.
3. Bakucs Z., Falkowski J., Ferto I.: Does Market Structure Influence Price Transmission in the Agro-food Sector? A Meta-analysis Perspective. Journal of Agriculture Economics, 65(1), 1–25 (2014).
4. Bayramoğlu A. T.: The Impact of Agricultural Commodity Price Increases on Agricultural Employment in Turkey. Procedia - Social and Behavioral Sciences, 143, 1058–1063 (2014).
5. Jaile-Benitez J. M., Ferrer-Comalat J. C., Linares-Mustarós S.: Determining the Influence Variables in the Pork Price, Based on Expert Systems. In: Gil-Aluja J., Terceño-Gómez A., Ferrer-Comalat J. C., Merigó-Lindahl J. M., Linares-Mustarós S. (eds.): Scientific Methods for the Treatment of Uncertainty in Social Sciences, 377, 81–92 (2015).
6. Patočková M.: Mléka je nadbytek, jeho ceny padají. In: IDNES.CZ. Available at http://ekonomika.idnes.cz/mleka-je-nadbytek-jeho-ceny-padaji-dli-/ekonomika.aspx?c=A080529_205803_ekonomika_dp, last accessed Nov 1, 2016.
7. Pechrová M., Medonos T.: Total factor productivity change in the Czech poultry sector. In: The 10th International Days of Statistics and Economics (MSED 2016). Slaný: Melandrium, 1806–1815 (2016).
8. Rumánková L.: Possibilities of Utilization of Univariate Time Series Analysis in Prices Modelling. In: Agrarian perspectives XXV. – Global and European Challenges for Food Production, Agribusiness and the Rural Economy. Prague: CULS Prague, 319–326 (2016).
9. Saengwong S., Jatuporn C., Roan S. W.: An Analysis of Taiwanese Livestock Prices: Empirical Time Series Approaches. Journal of Animal and Veterinary Advances 11(23), 4340–4346 (2012).
10. Šimpach O.: Statistical view of the current situation of Beekeeping in the Czech Republic. In: 6th International Days of Statistics and Economics. Slaný: Melandrium, 1054–1062 (2012).
11. Štolcová J., Homolka J.: Production Costs in the Fattening Period of Pigs and Assessment of the Influence of Selected Factors on the Amount of These Costs. Agris on-line Papers in Economics and Informatics, 4(4 – Special), 107–113 (2012).