COMPARISON OF LINEAR AND NONLINEAR MODELS FOR FORECASTING OF FOOD COMMODITY PRICES IN LATVIA

Oksana Pavlenko, Andrejs Matvejevs
Riga Technical University, Latvia
oksana.pavlenko@rtu.lv, andrejs.matvejevs@rtu.lv

Abstract. Analysing real-life data of commodity price dynamics is challenging, there can be non-stationary, non-linear, contain structural breaks. In this paper, we explore whether threshold models are preferable to linear autoregressive models (ARIMA) and whether the logistic smooth transition (LSTAR) model is preferable to the self-exciting threshold autoregressive (SETAR) model for important Latvian food commodity prices. Using historical prices of 16 most popular food products in Latvia over the last 18 years, we assess the goodness of fit of ARIMA (SARIMA), SETAR and LSTAR for each of the most popular commodity prices in Latvia and then compare the out-of-sample forecasts using measures RMSE and MAPE. Although different types of models appear to be most suitable for different commodities, even despite their similarity like fresh pork, chicken and beef, the overarching conclusion is that regime-switching models fit the prices of the majority of products better. ARIMA is the preferred model for some goods for construction of out-of-sample forecasts marginally more often than for the goodness of fit. Nevertheless, threshold models still appear superior in most cases. Additionally, we obtain rather large smoothness coefficients for most LSTAR models, which means that there are no significant reasons to prefer LSTAR to SETAR.

Keywords: time series, goodness of fit, out-of-sample forecast, structural break, ARIMA, SARIMA, LSTAR, SETAR.

Introduction

Forecasting food commodity prices is important because it helps producers plan their production and helps policymakers monitor food inflation, make informed decisions about economic policy, and ensure food security.

The Latvian food commodity prices have been volatile and unstable in recent years, and the data from 2005 till now may exhibit nonlinearity and structural breaks. Different events can lead to structural breaks in the data, such as changes in international food prices, shifts in trade policies, and economic shocks. For instance, the COVID-19 pandemic caused disruptions in global supply chains and led to an increase in food commodity prices in Latvia. Similarly, the 2014 Russian embargo on EU food products resulted in a drop in food exports from Latvia and a fall in some food commodity prices.

Various linear and nonlinear econometric models can be employed for forecasting commodity prices in different industries. Price dynamics may contain structural breaks, caused by events that differ for regions and particular countries. Forecasting with classical linear time series models (ARIMA, ARIMAX, SARIMA) often is not the best option. Threshold autoregressive models (SETAR) are considered to be more appropriate in the presence of structural breaks. Besides, switching of regimes may happen gradually rather than in an instant. This can be accounted for by smooth transition models like LSTAR (Logistic Smooth Transition Autoregressive Model).

There is no strong recommendation which type of a model is preferable for food commodity prices analysis and forecast. Researchers have used different models to forecast globally traded food commodity prices, including ARIMA and SARIMA models, Vector Autoregression (VAR) and Vector Error Correction (VEC) models and threshold models (SETAR and LSTAR). As it was shown by B.G.Hansen [1], threshold models like SETAR and LSTAR give rather good forecast comparing with a simple autoregression for the most frequently globally traded dairy commodity prices. V.P. de Albuquerquemello and others [2] show that the transition regime models for global corn prices give better forecast than the linear autoregressive models. M.A.Iquebal, Himardi Ghosh and Prajneshu [3] use three regime SETAR to Indian lac production data. Juho Valtiala [4] shows that the threshold autoregressive model fits the data of Finnish agricultural land prices better and improves the accuracy of price forecasts compared to the linear autoregression. STAR models are preferred in [5-7]. B.G.Hansen in [8] and Cathy W. S. Chen∗, Mike K. P. So and Feng-Chi Liu in [9] explain popularity of threshold models in economics and in finance.

Threshold models like SETAR and LSTAR are preferred over linear autoregression models when a series exhibits nonlinearity or structural breaks, as is often the case with economic and financial longitudinal data. These models allow for different regimes in the data, where the relationship between the variables may be different in each regime. This can lead to more accurate forecasts, especially when the data have nonlinear patterns or are subject to sudden changes due to unexpected events or shocks.

We analyse monthly prices for several mostly traded groups of products on the Latvian market: meet products (fresh pork, chicken, beef), bread (4 types), cereals (rice, buckwheat, oat flakes, wheat flour) and dairy products (milk, cottage cheese, butter, yoghurt). The analysis is based on publicly available monthly data from January 2005 to December 2022 published by the Central Statistical Bureau of Latvia [10]. First, we consider nominal prices only, then analyse the real prices.

The aim of the study is to understand whether using of threshold models instead of widely used ARIMA models can result in better goodness of fit and forecast. An additional goal is to establish which threshold model, SETAR or LSTAR, is superior.

Materials and methods

The idea of the autoregressive model (AR) is that the current value is explained by these series past values. The autoregressive moving average (ARMA) model is a more parsimonious model because it replaces large lags of the variable by errors in the previous steps. According to Box and Jenkins [11], it can be constructed for stationary time series only. In case of I(d) – nonstationary time series which becomes stationary after differencing d times – autoregressive integrated moving average model ARIMA(p,d,q) can be used [10]:

$$\Delta^d Y_t = a_0 + \sum_{i=1}^{p} a_i \Delta^d Y_{t-i} + \sum_{j-1}^{q} b_j \varepsilon_{t-j} + \varepsilon_t \, , \tag{1}$$

where Y_t – observed time series;

p – number of autoregressive terms;

q – number of moving average terms;

$\varepsilon_t \sim i.i.d.N(0,\sigma_\varepsilon^2)$ – white noise (estimation error in the moment t);

Δ – simple difference (lag) $\Delta Y_t = Y_{t-1}$. d usually 0 or 1, rarely 2.

This is a very popular linear time series model because of its simplicity. The parameters of a model can be estimated using the usual maximum likelihood method, implemented in most software. It can be improved including dependence of exogeneous variables (or their lags) and/or seasonality. Multiplicative seasonal ARIMA models (SARIMA(p,q,d)(P,Q,D)$_S$) combine usual ARIMA and seasonal ARIMA:

$$\Phi_P\left(\Delta^S\right)\phi(\Delta)\nabla_s^D \nabla^d Y_t = \delta + \Theta_q\left(\Delta^S\right)\theta(\Delta)\varepsilon_t \, , \tag{2}$$

where $\nabla_S{}^d = (1 - \Delta^S)^d$ – seasonal difference operator;

$\nabla^d = (1 - \Delta)^d$ – simple difference operator;

$\Phi_P(\Delta^S) = 1 - \Phi_1\Delta^S - \Phi_2\Delta^{2S} - \ldots - \Phi_P\Delta^{PS}$ – seasonal autoregressive operator;

$\Theta_Q(\Delta^S) = 1+\Theta_1\Delta^S + \Theta_2\Delta^{2S} + \ldots + \Theta_Q\Delta^{QS}$ – seasonal moving average operator;

$\phi(\Delta)$ – autoregressive operator;

$\theta(\Delta)$ – moving average operator [12];

s – number of seasons.

Threshold models are a very popular type of nonlinear time series models. It is a piecewise linear model with switching between different regimes, most often between two regimes. The self-exciting threshold AR with k-regimes (SETAR) [11] fits better than the linear ARIMA model if time series has some changes of a structure.

$$Y_t = \begin{cases} a_0^{(1)} + \sum_{i=1}^{p_1} a_i^{(1)} Y_{t-i} + \varepsilon_t^{(1)}, \text{if } X_{t-d} \le r_1 \\ \ldots \\ a_0^{(k)} + \sum_{i=1}^{p_k} a_i^{(k)} Y_{t-i} + \varepsilon_t^{(k)}, \text{if } r_{k-1} < X_{t-d} \le r_1 \end{cases} , \tag{3}$$

where p_1, …, p_k – orders of regression equations in corresponding regimes;

d – delay parameter;
$r_1, \ldots, r_k$ thresholds.

Besides, switching of regimes usually does not happen in an instant, but gradually in a smooth manner. This is implemented into smooth transition models like the Logistic Smooth Transition AutoRegressive model (LSTAR) and Exponential Smooth Transition AutoRegressive model (ESTAR). LSTAR is given by the following equation [11]

$$\begin{aligned} &Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + \theta\left[\beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p}\right] + \varepsilon_t, \\ &\theta = \left[1 + e^{-\gamma(Y_{t-1}-c)}\right]^{-1} \end{aligned} \tag{4}$$

where γ – smootheness parameter.

With increasing of γ the transition from

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_{p1} Y_{t-p} \text{ to } Y_t = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1) Y_{t-1} + \ldots + (\alpha_0 + \beta_0) Y_{t-p}$$

becomes sharper.

Having nonstationary, nonlinear data it is not recommended to take differences before fitting a threshold model [2], because classical unit root tests have low power in case of nonlinearity. Zivot-Andrews unit root test checks stationarity with breaks against a unit root in a time series. As nonlinearity may occur in many ways, there exist different tests for detecting nonlinearity. We use Ramsey Regression Error Specification test (RESET) [11] and Tsay test [13] of nonlinearity in time series against autoregression. The most popular measures of goodness of fit and of forecast accuracy are Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

Results and discussion

Nominal prices

First, we analyse the dynamics of food product nominal prices. As one can see from Fig.1 graphs, the series are rather different in each group of similar products except for bread. However, for most of them we see a change of structure in the beginning of 2009 and a sharp increase in 2021 and 2022.

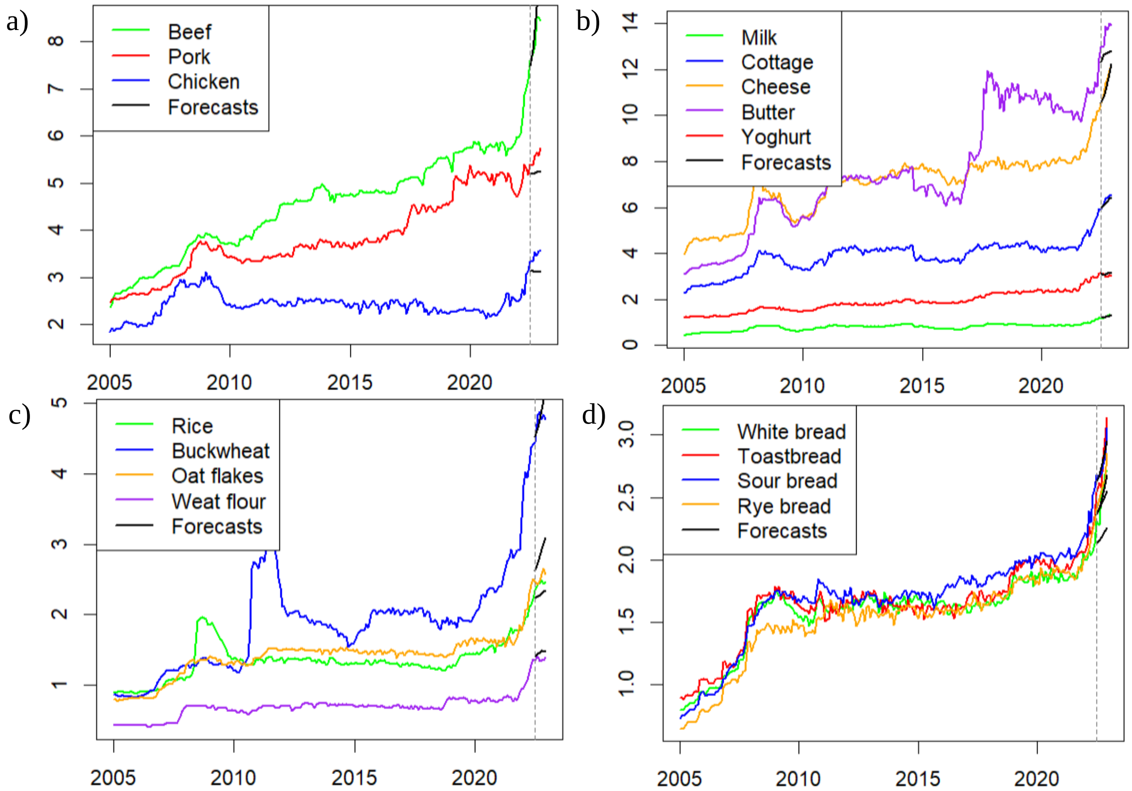

Fig. 1. **Graphical representation of the data (nominal prices in EUR·kg^{-1}) and forecast of the last 6 values by the best model**

According to the Dickey-Fuller Augmented test (unit root test) all the data appeared to be nonstationary, mostly I(1), some even I(2) (Sour bread and Toast bread). The Zivot-Andrews Unit Root test pointed potential breaks close to the end of the period of data. The RESET test shows strong nonlinearity for almost all dairy products, for all types of bread, for beef and oat flakes. For the others (fresh pork, chicken, yoghurt, rice, buckwheat and wheat flour) linearity was not rejected by this test. But the Tsay test additionally rejects linearity for chicken and all cereal products. Only for pork and yoghurt the price linearity is not rejected. Despite that, we built all three types of models (ARIMA (or SARIMA), LSTAR and SETAR) for each product, choosing the most appropriate model of each type for each product. We compared their goodness of fit using MAPE and RMSE, and then, shortening the training period by leaving the last six months of the observation period for the testing period, reestimated the models on the restricted dataset, built the forecasts and compared using RMSE and MAPE. The results are seen in Tables 1 to 4. The best model type with its MAPE and RMSE is shown in bold for each product.

Table 1

Analysis result for meat products

Model	**Beef**	**Pork**	**Chicken**
Best model by *auto.arima*	ARIMA(4,1,0) with drift	ARIMA(0,1,0) with drift	ARIMA(2,0,2) with non-zero mean
MAPE for goodness of fit	1.157% (ARIMA) 1.115% (SETAR) **1.031% (LSTAR)**	1.478% (ARIMA) **1.422% (SETAR)** 1.435% (LSTAR)	2.376% (ARIMA) 2.341% (SETAR) **2.316% (LSTAR)**
RMSE for goodness of fit	0.0889 (ARIMA) 0.0859 (SETAR) **0.0725 (LSTAR)**	0.0868 (ARIMA) **0.0835 (SETAR)** 0.0844 (LSTAR)	0.0830 (ARIMA) 0.0833 (SETAR) **0.0818 (LSTAR)**
MAPE for forecast	9.492% (ARIMA) **6.120% (SETAR)** 14.537% (LSTAR)	**5.886% (ARIMA)** 7.524% (SETAR) 7.523% (LSTAR)	10.901% (ARIMA) 11.971% (SETAR) **9.805% (LSTAR)**
RMSE for forecast	0.8528 (ARIMA) **0.7256 (SETAR)** 1.6413 (LSTAR)	**0.3480 (ARIMA)** 0.4447 (SETAR) 0.4447 (LSTAR)	0.3979 (ARIMA) 0.4404 (SETAR) **0.3576 (LSTAR)**

Table 2

Analysis result for dairy products

Model	**Milk**	**Cottage**	**Cheese**	**Butter**	**Yoghurt**
Best by *auto.arima*	ARIMA (1,1,2)	ARIMA (1,1,2)	ARIMA(1,1,2) with drift	ARIMA (3,1,1) with drift	ARIMA (2,1,2) with drift
MAPE for goodness of fit	2.068% (ARIMA) 2.136% (SETAR) **2.052% (LSTAR)**	2.183% (ARIMA) **2.127% (SETAR)** 2.148% (LSTAR)	1.882% (ARIMA) **1.715% (SETAR)** 1.912% (LSTAR)	2.521% (ARIMA) 2.487% (SETAR) **2.400% (LSTAR)**	1.765% (ARIMA) 1.827% (SETAR) **1.706% (LSTAR)**
RMSE for goodness of fit	0.0221 (ARIMA) 0.0224 (SETAR) **0.0219 (LSTAR)**	0.1174 (ARIMA) **0.1104 (SETAR)** 0.1152 (LSTAR)	0.1778 (ARIMA) **0.1623 (SETAR)** 0.1751 (LSTAR)	0.3033 (ARIMA) 0.3074 (SETAR) **0.2879 (LSTAR)**	**0.0501 (ARIMA)** 0.0507 (SETAR) 0.0503 (LSTAR)
MAPE for forecast	**2.070%(ARIMA)** 6.839% (SETAR) 5.928% (LSTAR)	**1.799% (ARIMA)** 1.825%(SETAR) 5.986% (LSTAR)	8.763% (ARIMA) 1.367% (SETAR) **1.253% (LSTAR)**	**6.695% (ARIMA)** 16.770% (SETAR) 15.722% (LSTAR)	**2.814%(ARIMA)** 7.928% (SETAR) 5.221% (LSTAR)
RMSE for forecast	**0.0304 (ARIMA)** 0.1034 (SETAR) 0.0887 (LSTAR)	**0.1273 (ARIMA)** 1.7126 (SETAR) 0.4301 (LSTAR)	1.1502 (ARIMA) 0.2234 (SETAR) **0.1772 (LSTAR)**	**0.9689 (ARIMA)** 2.4087 (SETAR) 2.3011 (LSTAR)	**0.0966 (ARIMA)** 0.2712 (SETAR) 0.1837 (LSTAR)

Table 3

Analysis result for cereal products

Model	**Rice**	**Buckwheat**	**Oat flakes**	**Wheat flour**
Best model by *auto.arima*	ARIMA (1,1,1)	SARIMA (1,1,0)(0,0,1)[12] with drift	SARIMA(1,1,2)(0,0,2)[12]	SARIMA (2,1,0)(0,0,1)[12] with drift
MAPE for goodness of fit	2.0615% (ARIMA) **1.945% (SETAR)** 2.201% (LSTAR)	**2.792% (ARIMA)** 3.084% (SETAR) 3.002% (LSTAR)	2.029% (ARIMA) 1.982% (SETAR) **1.964% (LSTAR)**	2.507%(ARIMA) 2.499% (SETAR) **2.385% (LSTAR)**
RMSE for goodness of fit	0.04530 (ARIMA) 0.0444 (SETAR) **0.0441 (LSTAR)**	0.0984 (ARIMA) 0.0991 (SETAR) **0.0904 (LSTAR)**	0.0433 (ARIMA) 0.0422 (SETAR) **0.0400 (LSTAR)**	0.0263 (ARIMA) 0.0264 (SETAR) **0.0255 (LSTAR)**
MAPE for forecast	6.270% (ARIMA) **5.960% (SETAR)** 6.650% (LSTAR)	5.543% (ARIMA) 5.0345% (SETAR) **3.986% (LSTAR)**	**13.470% (ARIMA)** 26.256% (SETAR) 20.308% (LSTAR)	**6.989% (ARIMA)** 37.384% (SETAR) 12.373% (LSTAR)
RMSE for forecast	0.1563 (ARIMA) **0.1492 (SETAR)** 0.1659 (LSTAR)	0.2974 (ARIMA) 0.3019 (SETAR) 0.2221 **(LSTAR)**	**0.3581 (ARIMA)** 0.7588 (SETAR) 0.5592 (LSTAR)	**0.1005 (ARIMA)** 0.5957 (SETAR) 0.1776 (LSTAR)

Table 4

Analysis result for bread products

Model	**White bread**	**Toast bread**	**Sour bread**	**Rye bread**
Best model by *auto.arima*	ARIMA(2,1,2)	ARIMA(3,2,1)	ARIMA(1,1,2) with drift	ARIMA(1,1,2) with drift
MAPE for goodness of fit	2.084% (ARIMA) 2.079% (SETAR) **2.059% (LSTAR)**	2.339% (ARIMA) 2332% (SETAR) **2.239% (LSTAR)**	1.801% (ARIMA) 1.790% (SETAR) **1.771% (LSTAR)**	2.263% (ARIMA) 2.230% (SETAR) **2.158% (LSTAR)**
RMSE for goodness of fit	0.0469 (ARIMA) 0.0475 (SETAR) **0.0432 (LSTAR)**	0.0533 (ARIMA) 0.0538 (SETAR) **0.0485 (LSTAR)**	0.0434 (ARIMA) 0.0421 (SETAR) **0.0412 (LSTAR)**	0.0458 (ARIMA) 0.0465 (SETAR) **0.0427 (LSTAR)**
MAPE for forecast	15.092% (ARIMA) 14.470% (SETAR) **12.375% (LSTAR)**	15.672% (ARIMA) 11.468% (SETAR) **11.019% (LSTAR)**	3.239% (ARIMA) 1.920**% (SETAR)** 6.996% (LSTAR)	10.609% (ARIMA) 4.475% (SETAR) **3.175% (LSTAR)**
RMSE for forecast	0.4245 (ARIMA) 0.4078 (SETAR) **0.3444 (LSTAR)**	0.5006 (ARIMA) 0.3717 (SETAR) **0.3570 (LSTAR)**	0.1299 (ARIMA) **0.0608 (SETAR)** 0.2380 (LSTAR)	0.3077 (ARIMA) 0.1317 (SETAR) **0.0940 (LSTAR)**

For all analysed products, nonlinear models (LSTAR or SETAR) give better goodness of fit by at least one criterion, mostly by both, even for those whose linearity was not rejected. Only for yoghurt, ARIMA was chosen as the best according to RMSE and for buckwheat according to MAPE. However, comparing out-of-sample forecasts, only in group of bread all series performed better by nonlinear models, mostly LSTAR, but in other groups better forecasts more often (but not for all series in each group) are given by ARIMA (or SARIMA) models. Multiple seasonal ARIMA models were chosen as best ARIMA only for a part of cereal products, then they gave the best forecast only for oat flakes and wheat flour.

Predicting the nominal prices, LSTAR was chosen more often than SETAR. However, the value of a smoothness parameter in the chosen as best models LSTAR mostly was rather large, what pointed to the closeness to SETAR model. The smoothness parameter was not large only in the model for white bread (16.23).

Real prices

Comparing Fig.1 with Fig.2, we see that for most commodities their price dynamics resembles the Consumer Price Index of Latvia for the same period [14], especially higher growth before 2008 and after 2021, as well as a fall in 2010.

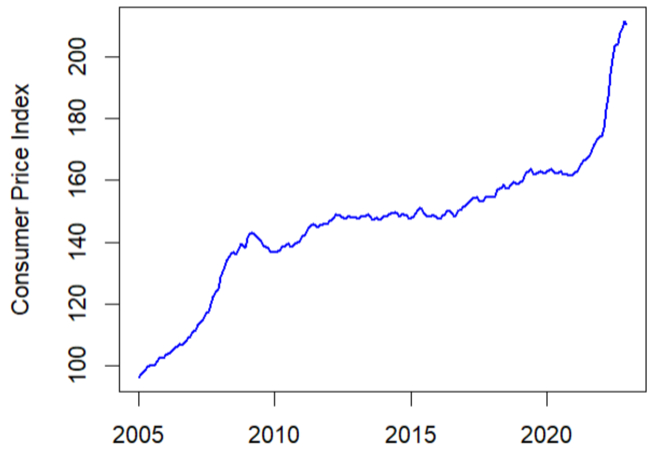

Fig. 2. **Consumer Price Index of Latvia (average in 2005 as 100)**

This means that CPI should be definitely included into the models. One option is to include it as an exogeneous variable into each model. However, it is not a good solution for nonlinear models.

The other way to account for inflation is to analyse the real prices of the products, i.e. the nominal prices divided by the Consumer Price Index (average in 2005 as 1). For the visual presentation, we standardize the series, taking each series' first value (in January 2005) as 1.

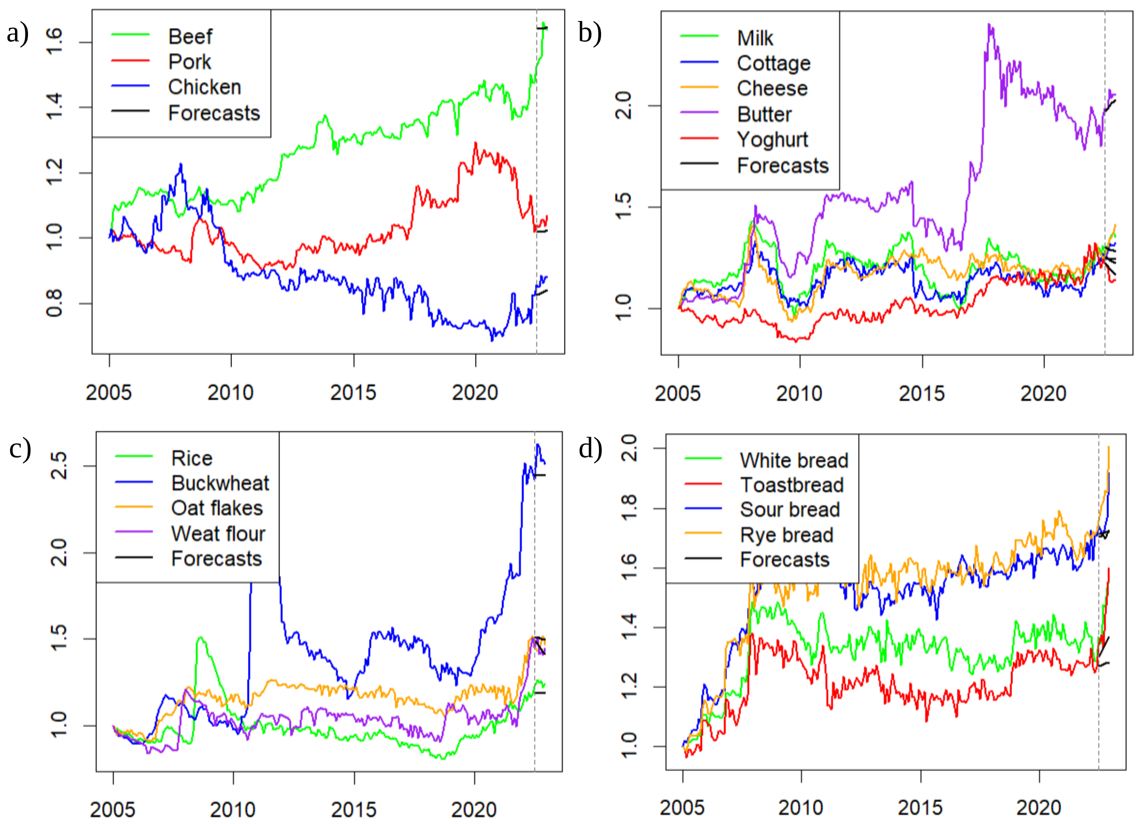

Fig. 3. **Graphical representation of the data (real standardized prices) and forecast of the last 6 values by the best model**

Completely all series of real prices were identified as I(1) by ADF test and almost all automatically chosen in R program ARIMA models appeared integrated of the first order except of cottage where the preference is given to ARMA(2,2) and for milk chosen AR(1) which is not full enough. Seasonality appeared in the models for three out of four types of bread that led to multiple seasonal ARIMA models. The correction with CPI lead to the fact that for less data linearity is rejected by nonlinearity tests, especially by RESET test. Now only butter, buckwheat and toast bread data were declared nonlinear by both tests. So, we have chosen the most appropriate model of each type for each product and compared them again. The threshold models, SETAR or LSTAR, are found to be better for the goodness of fit of all analysed series. Also, nonlinear models give better out-of-sample forecast more often comparing with ARIMA, but not so often as goodness of fit. ARIMA gives better forecast for cottage cheese, rice, buckwheat, sour bread and rye bread both by RMSE and MAPE. (See Tables 5,6,7,8 for the results. Best models are shown in bold.)

Table 5

Analysis result for the real prices of meat products

Model	**Beef**	**Pork**	**Chicken**
Best model by *auto.arima*	ARIMA(2,1,0) with drift	ARIMA(0,1,0)	ARIMA(0,1,1)
MAPE for goodness of fit	1.111% (ARIMA) 1.098% (SETAR) **1.091% (LSTAR)**	1.503% (ARIMA) **1.450% (SETAR)** 1.468% (LSTAR)	2.282% (ARIMA) 2.351%(SETAR) **2.206% (LSTAR)**
RMSE for goodness of fit	0.0214 (ARIMA) **0.0203 (SETAR)** 0.0207 (LSTAR)	0.0220 (ARIMA) **0.0213 (SETAR)** 0.0217 (LSTAR)	0.0280 (ARIMA) 0.0280 (SETAR) **0.0275 (LSTAR)**
MAPE for forecast	6.836 % (ARIMA) **3.405% (SETAR)** 6.993% (LSTAR)	2.664% (ARIMA) 3.654% (SETAR) **2.627% (LSTAR)**	6.029% (ARIMA) **4.208% (SETAR)** 5.188% (LSTAR)
RMSE for forecast	0.1216 (ARIMA) **0.0698 (SETAR)** 0.1242 (LSTAR)	0.0312 (ARIMA) 0.0420 (SETAR) **0.0308 (LSTAR)**	0.0545 (ARIMA) **0.0384 (SETAR)** 0.0473 (LSTAR)

Table 6

Analysis result for the real prices of dairy products

Model	**Milk**	**Cottage**	**Cheese**	**Butter**	**Yoghurt**
Best model by *auto.arima*	ARIMA(1,0,0) with non-zero mean	ARIMA(2,0,2) with non-zero mean	ARIMA (2,1,0)	ARIMA (2,1,2)	ARIMA (2,1,2)
MAPE for goodness of fit	2.042% (ARIMA) **2.005%(SETAR)** 2.051% (LSTAR)	2.087% (ARIMA) **2.027%(SETAR)** 2.067% (LSTAR)	1.865% (ARIMA) 1.739% (SETAR) **1.738% (LSTAR)**	2.446% (ARIMA) 2.356% (SETAR) **2.355% (LSTAR)**	1.725% (ARIMA) **1.705% (SETAR)** 1.716% (LSTAR)
RMSE for goodness of fit	2.0416 (ARIMA) **2.005 (SETAR)** 2.051 (LSTAR)	0.0311 (ARIMA) **0.0310 (SETAR)** 0.0314 (LSTAR)	0.0286 (ARIMA) **0.0267 (SETAR)** 0.0269 (LSTAR)	0.0602 (ARIMA) **0.0564 (SETAR)** 0.0588 (LSTAR)	0.02448 (ARIMA) 0.02448 (SETAR) **0.02441 (LSTAR)**
MAPE for forecast	4.132% (ARIMA) **4.089%(SETAR)** 4.636% (LSTAR)	**4.978%(ARIMA)** 6.371% (SETAR) 6.329% (LSTAR)	**7.221%(ARIMA)** 7.715% (SETAR) 7.854% (LSTAR)	5.383% (ARIMA) **1.540% (SETAR)** 1.684% (LSTAR)	5.703% (ARIMA) **2.808% (SETAR)** 4.594% (LSTAR)
RMSE for forecast	0.0628(ARIMA) **0.0617 (SETAR)** 0.0700 (LSTAR)	**0.0718 (ARIMA)** 0.0930 (SETAR) 0.0925 (LSTAR)	**0.1120 (ARIMA)** 0.1167 (SETAR) 0.1200 (LSTAR)	0.1140 (ARIMA) **0.0399 (SETAR)** 0.0442 (LSTAR)	0.0735 (ARIMA) **0.0383 (SETAR)** 0.0607 (LSTAR)

Table 7

Analysis result for the real prices of cereal products

Model	Rice	Buckwheat	Oat flakes	Wheat flour
Best model by *auto.arima*	ARIMA(1,1,0)	ARIMA(1,1,0)	ARIMA(1,1,0)	ARIMA(0,1,0)
MAPE for goodness of fit	2.109% (ARIMA) 1.941% (SETAR) **1.925% (LSTAR)**	2.741% (ARIMA) **2.725% (SETAR)** 2.796% (LSTAR)	1.939% (ARIMA) **1.874% (SETAR)** 1.883% (LSTAR)	2.496% (ARIMA) **2.397% (SETAR)** 2.534% (LSTAR)
RMSE for goodness of fit	0.0325 (ARIMA) 0.0308 (SETAR) **0.0305 (LSTAR)**	0.0725 (ARIMA) 0.0712 (SETAR) **0.0685 (LSTAR)**	0.0313 (ARIMA) 0.0306 (SETAR) **0.0305 (LSTAR)**	0.0365 (ARIMA) **0.0348 (SETAR)** 0.0355 (LSTAR)
MAPE for forecast	**4.524% (ARIMA)** 6.959% (SETAR) 5.446% (LSTAR)	**3.985% (ARIMA)** 9.163% (SETAR) 5.279% (LSTAR)	3.289% (ARIMA) 3.144% (SETAR) **2.862% (LSTAR)**	4.616% (ARIMA) **2.186% (SETAR)** 3.334% (LSTAR)
RMSE for forecast	**0.0589(ARIMA)** 0.0898 (SETAR) 0.0701(LSTAR)	**0.1159 (ARIMA)** 0.2562 (SETAR) 0.1329 (LSTAR)	0.05556 (ARIMA) 0.0529 (SETAR) **0.0483 (LSTAR)**	0.0731 (ARIMA) **0.0332 (SETAR)** 0.0521 (LSTAR)

Table 8

Analysis result for bread products

Model	White bread	Toast bread	Sour bread	Rye bread
Best model by *auto.arima*	ARIMA(1,1,0)	SARIMA(0,1,2)(1,0,1)[12]	SARIMA(0,1,1)(1,0,2)[12]with drift	SARIMA(0,1,1)(1,0,1)[12]
MAPE for goodness of fit	2.010% (ARIMA) 2.001% (SETAR) **1.928% (LSTAR)**	2.287% (ARIMA) 2.246% (SETAR) **2.204% (LSTAR)**	1.709% (ARIMA) 1.684% (SETAR) 1.665**% (LSTAR)**	2.115% (ARIMA) **2.032% (SETAR)** 2.077% (LSTAR)
RMSE for goodness of fit	0.0367 (ARIMA) 0.0351 (SETAR) **0.0348 (LSTAR)**	0.0384 (ARIMA) 0.0373 (SETAR) **0.0366 (LSTAR)**	0.0349 (ARIMA) 0.0349 (SETAR) **0.0346 (LSTAR)**	0.0450 (ARIMA) **0.0421 (SETAR)** 0.0432 (LSTAR)
MAPE for forecast	11.370% (ARIMA) **8.268% (SETAR)** 9.144% (LSTAR)	10.833% (ARIMA) **10.521% (SETAR)** 11.754% (LSTAR)	**3.158% (ARIMA)** 4.213% (SETAR) 3.577% (LSTAR)	**7.049% (ARIMA)** 8.374% (SETAR) 7.930% (LSTAR)
RMSE for forecast	0.1867 (ARIMA) **0.1362 (SETAR)** 0.1524 (LSTAR)	0.1875 (ARIMA) **0.1819 (SETAR)** 0.2023 (LSTAR)	**0.0855 (ARIMA)** 0.1083 (SETAR) 0.0972 (LSTAR)	**0.1530 (ARIMA)** 0.1808 (SETAR) 0.1723 (LSTAR)

Predicting the real prices, SETAR was chosen more often than LSTAR. However, even if LSTAR was preferred, the value of the smoothness parameter was rather large, close to 100, what pointed to the closeness to the SETAR model.

Conclusions

The aim of the study to compare threshold models to linear ARIMA models and LSTAR to SETAR for key Latvian food commodity prices has been achieved.

Completely all analysed data of commodity prices in Latvia appeared to be nonstationary in the observed period. Nonlinearity is indicated for most of them. After turning to real prices, nonlinearity tests did not reject linearity of some series. This implies that deflating the data and working with real price values has a potential to reduce nonlinearity.

The obtained results let us conclude that the threshold models produce better forecast more often than autoregression.

Regime switching models, SETAR and LSTAR, give better goodness of fit of commodity prices than linear ARIMA models, even for some series whose linearity is not rejected. The average improvement in MAPE when using the best threshold model (if chosen) instead of ARIMA is 1.834% for real prices (3.330% for nominal prices).

Comparing the two threshold model types, SETAR was chosen more often than LSTAR for predicting the real prices, despite LSTAR being more intuitively suitable with its smooth transition. The chosen LSTAR models have a large smoothness coefficient, which means that smoothness in the estimated threshold models is not important. Therefore, there are no significant reasons to prefer LSTAR to SETAR for prediction of food commodity prices in Latvia.

The analysis performed in this study demonstrates that real-life series of prices or other economic data often contain breaks, non-linearities and other peculiarities that require going beyond common time-series methods such as linear ARIMA models in order to produce reliable forecasts. At the same time, it is also important to stick to the principle of parsimony when choosing the most appropriate model, since introduction of additional parameters such as transition smoothness in LSTAR might not result in smaller errors.

Author contributions

Formal analysis, O.P.; investigation, O.P, data curation, O.P and A.M..; writing – original draft preparation, O.P; visualization, O.P.; validation, A.M. All authors have read and agreed to the published version of the manuscript.

References

[1] Hansen G. Forecasting prices of dairy commodities – a comparison of linear and nonlinear models. Irish Journal of Agricultural and Food Research. 59(1) 2020 DOI: 10.15212

[2] Albuquerquemello V. P., Medeiros R. K., Jesus D. P., Oliveira, F. A. The role of regime transition models for corn prices forecasting. Revista de Economia e Sociologia Rural, 60(2), 2022, e236922. DOI: 10.1590/1806-9479.2021.236922

[3] Iquebal M.A., Ghosh H., Prajneshu. Fitting of SETAR Three-regime nonlinear time series model to Indian lac production data through genetic algorithm. Indian Journal of AgriculturalSciences 83 (12): 1406-8, December 2013.

[4] Valtiala J. Testing for regime-switching behaviour in Finnish agricultural land prices. Agricultural Finance Review, Vol. 81 No. 2, 2021, pp. 292-305. DOI: 10.1108/AFR-03-2020-0037.

[5] Zeng J.H., Chang C.P., Lee C.C. Are Fruit and Vegetable Prices Non-linear Stationary? Evidence from Smooth Transition Autoregressive Models. Economics Bulletin, AccessEcon, vol. 31(1), 2011, pp. 189-207.

[6] Aliyev F. Testing Market Efficiency with Nonlinear Methods: Evidence from Borsa Istanbul. Int. J. Financial Stud. 2019, 7(2), 27; DOI: 10.3390/ijfs7020027

[7] Djeddour K., Boularouk Y. Application of Threshold Autoregressive Model: Modeling and Forecasting Using U.S. Exports of Crude Oil Data. American Journal of Oil and Chemical Technologies 1(9), November 2013, pp. 2326-6570. DOI: 10.14266/ajoct19-1.

[8] Hansen, G. Threshold autoregression in economics. Statistics and Its Interface. Volume 4 (2011) pp. 123-127.

[9] Cathy W. S. Chen, Mike K. P. So and Feng-Chi Liu. A review of threshold time series models in Finance. Statistics and Its Interface Volume 4, 2011, pp. 167-181.

[10] Atsevišķu preču un pakalpojumu vidējās mazumtirdzniecības cenas (eiro par 1kg, ja citādāk – norādīts īpaši) – Preces un pakalpojumi un Laika periods (Average retail prices of individual goods and services (euro per 1kg, if otherwise - specified separately) - Goods and services and Time period) [online][11.02.2023] Available at: https://data.stat.gov.lv/pxweb/lv/OSP_PUB/START_VEK_PC_PCC/PCC010m/ (In Latvian)

[11] Enders, W. Applied econometric time series Fourth edition, Publisher: Wiley; 4 edition (October 22, 2014) ISBN-13: 978-1118808566 ISBN-10: 1118808568

[12] Shumway, R. H., Stoffer, D. S. Time series analysis and its applications With R Examples, Third or Fourth edition Springer Science + Business Media, LLC 2011, 2017.

[13] Tsay, R.S. Nonlinearity tests for time series. Biometrica 73: 461-466, 1986.

[14] Patēriņa cenu indeksi un pārmaiņas grupās un apakšgrupās (ECOICOP) 1996M01 - 2023M04 (Consumer price indices and changes in groups and subgroups (ECOICOP) 1996M01 - 2023M04) [online][11.02.2023] Available at: https://data.stat.gov.lv/pxweb/lv/OSP_PUB/START__VEK__PC__PCI/PCI020m/ (In Latvian)