

Forecasting Wheat Prices Based on Past Behavior: Comparison of Different Modelling Approaches

Joana Dias[1,2(✉)] and Humberto Rocha[1,2]

[1] CeBER and Faculty of Economics, University of Coimbra, Coimbra, Portugal
joana@fe.uc.pt, hrocha@mat.uc.pt
[2] INESC-Coimbra, University of Coimbra, Coimbra, Portugal

Abstract. Being able to accurately forecast the evolution of wheat prices can be a valuable tool. Most of the published works apply classical forecasting models to wheat price time series, and they do not always perform out-of-sample testing. This work compares five modelling approaches for wheat price forecasts, using only past values of the time series. The models performance is assessed considering out-of-sample data only, by considering a sliding and growing time window that will define the data used to determine the models parameters, and the data used for out-of-sample forecasts.

Keywords: Agriculture · Machine learning · Price forecasting · Wheat

1 Introduction

According to the Food and Agriculture Organization of the United Nations (FAO), wheat is the most important grain source for humans, and it is grown in more land than any other commercial crop [1]. The demand has also increased due to its use as biofuels.

Wheat prices can be influenced by many different factors: competition in the main international wheat markets, existing regulations and policies, weather conditions, among other factors. Being able to develop models capable of accurately modelling and being able to explain wheat prices evolution is, thus, a difficult challenge. The United States (US) farm programs or the European Union Common Agricultural Policy are examples of policy factors influencing wheat production and price [2]. Brunner [3], examines the historical effects of El Niño on world primary commodity prices, reaching the conclusion that it has statistically significant economic effects. Hill et al. [4] evaluate the effect of using seasonal climate forecasts on international wheat economy. Ubilava [5] also follows this trend, reaching the conclusion that El Niño influences wheat prices.

In [6] the authors present an asymmetric-error generalized autoregressive conditional heteroscedasticity (GARCH) model, applying it to forecasts of U.S. soybean, sorghum and wheat prices. GARCH models can capture the phenomenon of volatility clustering, characterized by periods where the dependent variable presents high

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11621, pp. 167–182, 2019.
https://doi.org/10.1007/978-3-030-24302-9_13

volatility, and others where the time series are almost constant. The authors use wheat prices from 1913 to 2000, adjusted for inflation.

Benavides [7] studies the volatility accuracy of volatility forecast models for wheat future prices returns, applying GARCH and ARCH models, and also an option implied and a composite forecast model. The authors recommend the use of the composite forecast model if both historical values and option implied volatility forecasts are available.

Yang et al. [2] study wheat future prices and volatility transmission for the United States, Canada and European Union in the period 1996 to 2002. They observe that Canadian prices are more influenced by the US prices than the other way around. Europe, on the other hand, appears to be self-dependent, having some influence on the US prices on the long run. The authors reach the conclusion that none of the three markets can be identified as being the international market leader.

In [8] the authors study the relationships among wheat prices from five different countries in the period 1981–1999. The objective is to discover dynamic causal relationships among these prices. Contradicting previous results, the authors reach the conclusion that US and Canada are leaders in these markets, with US having significant effects on the three markets other than Canada.

Roche and McQuinn [9] consider the Irish grain market and the influence of the British grain prices on Irish prices. The authors use a multivariate GARCH model, where relative effects of past variances and covariances are determined entirely by the data and are not decided by the user.

A season-average futures price forecasting model for corn, soybeans and wheat is presented in [10]. This model considers future prices, farm prices and marketing weights.

Jumah and Kunst [11] consider barley and wheat prices and study the use of seasonal models, since these grains are subject to seasonal variation due to the biological growth of the plant, related also to climatic factors. The authors observe that prices tend to increase in the first and fourth quarter of the year, and to fall in the third, although seasonal cycles are different for different countries.

Arshad and Hameed [12] investigate the relationships between cereal prices and petroleum prices. Data from 1980 to 2008 was considered, and the authors reached evidence of a long-run equilibrium relationship between the two product prices.

Algieri [13] develops a vector error correction model considering a broad range of explanatory variables: market-specific variables, macroeconomic variables, financial factors and weather conditions. The author considers data within the period 1980–2012. The author states that changes observed in wheat prices are like a roller coaster, and seem to be inconsistent with supply and demand fundamentals. Wheat price movements can be explained looking at speculation, global demand and real effective exchange rate. Speculation in the futures market is one of the reasons justifying wider price changes. Chen et al. [14] had already considered the effect of exchange rates in commodity prices. In [15] it is also possible to find the development of a structural vector autoregression model for wheat price variation, considering four structural factors: global real economic activity and commodity demand; wheat-specific supply and demand factors; speculative or precautionary demand; financial speculation, commodity index trading and comovement. The authors conclude that wheat price spikes can be mostly explained by shocks to current supply.

Ahumada and Cornejo [16] focus on three cereals: corn, soybeans and wheat. They try to improve the accuracy of price forecast models by explicitly using price cross-dependence among these products, considering equilibrium correction models.

There are not many examples in the literature of machine learning approaches applied to wheat price forecasting. Khamis and Abdullah [17] investigate the use of backpropagation neural networks and nonlinear autoregressive models with exogenous inputs networks to estimate the price of wheat using as inputs the prices of other three grains: oats, barley and soybeans. Historical values from 1978 to 2012 were used, and the authors concluded that the latter model performed better.

This paper presents a different approach for wheat price forecasting. The main goal is to be able to predict the monthly wheat prices for the next six months period using information on past prices only. A rolling window is considered, where prices that become known are incorporated into the training of the forecasting model, and new predictions are made for the next six months. Five different methods are tested: Autoregressive Integrated Moving Average (ARIMA) models, Classification and Regression Trees (CART), Random Forests (RF), Support Vector Machines (SVM) and Multivariate Adaptive Regression Splines (MARS). All the models are tested on out-of-sample data.

Section 2 describes the data used and an exploratory analysis is performed. The models that are used are also described. Section 3 presents the main computational results. Section 4 presents the main conclusions and paths for future research.

2 Materials and Methods

2.1 Data

According to the United States Wheat Associates, there are six classes of wheat grown in the United States that differ in color, hardness and growing season. In this work we consider the export prices of hard red winter wheat (accounting for about 40% of the total wheat production), delivered at the United States Gulf port for prompt or thirty days shipment. One of the decisions that has to be made is whether real prices or nominal prices should be modelled and predicted. Real prices (also known as constant prices) are indexed to a given year. All values are "deflated", meaning that the effect of the inflation with respect to that year is removed. Nominal prices (also known as current prices) include the effect of inflation. In this paper we choose to work with nominal prices. Monthly prices from February 1992 to February 2017 are considered. Figure 1 shows the price evolution during this time period. Looking at Fig. 1, it is not possible to clearly identify trends, cycles or seasonality. This is also the case looking at Fig. 2, where the three components of the time series are depicted: seasonal component, trend-cycle component, remainder component, considering an additive model based on [18] (STL decomposition). Figure 3 presents the same data but with all values of the same month plotted together. The horizontal lines indicate the mean value for each month. The plot is not particularly revealing, since there are not many differences for the different months.

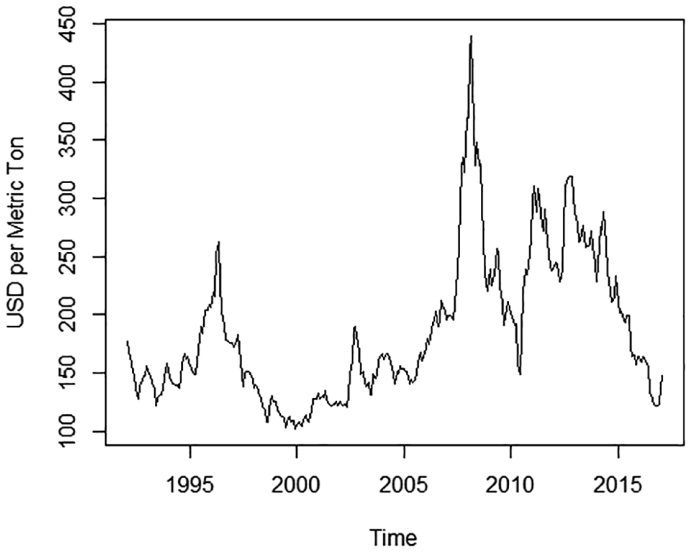


Fig. 1. Evolution of hard red winter wheat prices

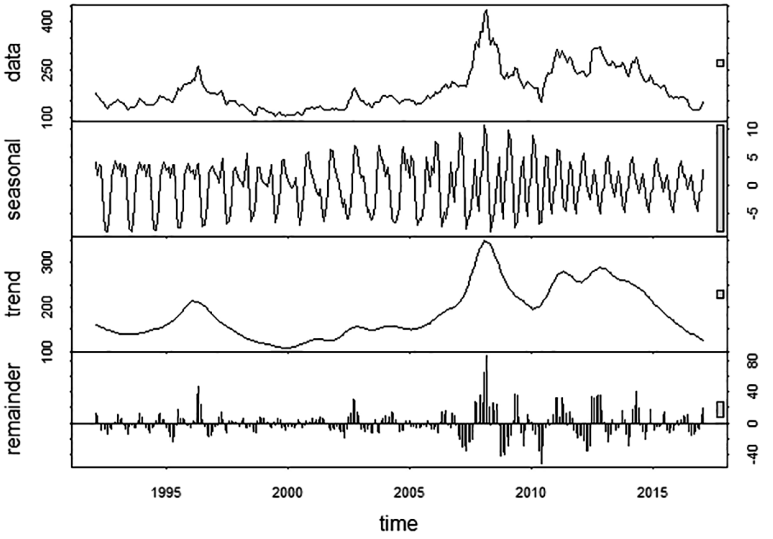


Fig. 2. STL decomposition

This work investigates the possibility of using forecasting models to predict monthly prices considering past information on prices only. It is thus important to understand whether there is a linear relationship between lagged values of this time series. One way of visually detecting autocorrelations is by looking at the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. Defining

as y_t the value of the time series at time period t and y_{t-k} the value of the time series k periods before, the ACF plot depicts the autocorrelations between y_t and y_{t-k} for different values for the lag k (Fig. 4). As can be observed, the ACF declines very slowly. Up to 24 lags are statistically significantly different from zero.

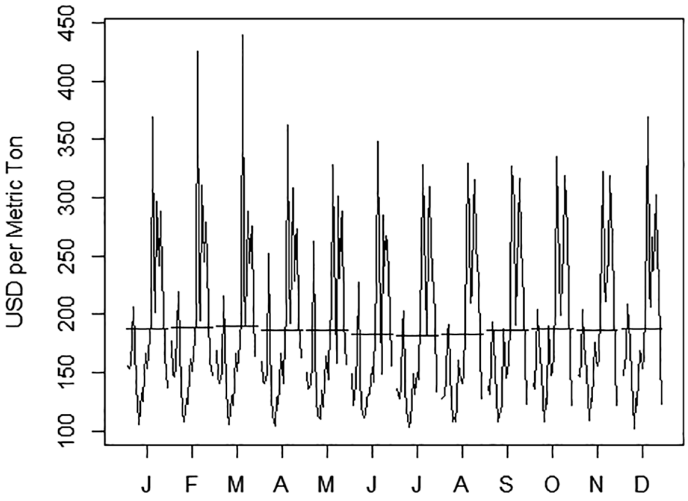


Fig. 3. Seasonal plot of monthly prices

The PACF measures the relationship between y_t and y_{t-k} after removing the effects of other time lags (from 1 to k − 1). Figure 5 depicts the PACF plot. After the first lag, the PACF drops dramatically. Most PACFs after lag 2 are statistically insignificant. A stationary time series is such that its properties will not depend on the time of the observation. When a series is not stationary, it is possible to make it stationary by differencing. One way to determine if differencing can be beneficial is to use unit root tests. One of such tests is the Augmented Dickey-Fuller Test (ADF). The null-hypothesis for an ADF test is that the data are non-stationary. So large p-values are indicative of non-stationarity, and small p-values suggest stationarity. Using the usual 5% threshold, differencing is required if the p-value is greater than 0.05. Applying ADF test to this time series, the p-value is 0.14. Another popular unit root test is the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. This reverses the hypotheses, so the null-hypothesis is that the data are stationary. In this case, small p-values (e.g., less than 0.05) suggest that differencing is required. In this case the p-value obtained is equal to 0.01. Figure 6 shows the series after first order differencing. This series is now stationary as can be concluded by looking at the ADF and KPSS tests, that present p values equal to 0.01 and 0.10 respectively.

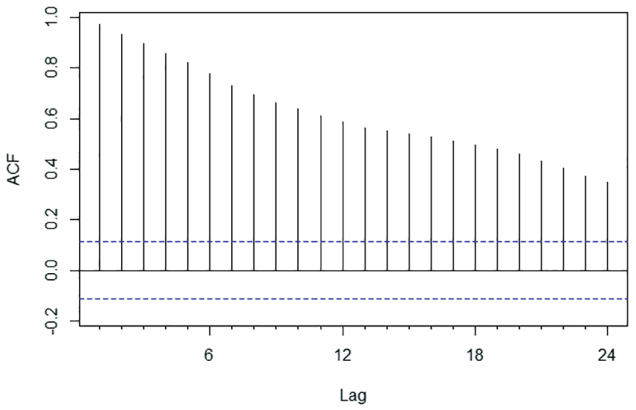


Fig. 4. ACF plot

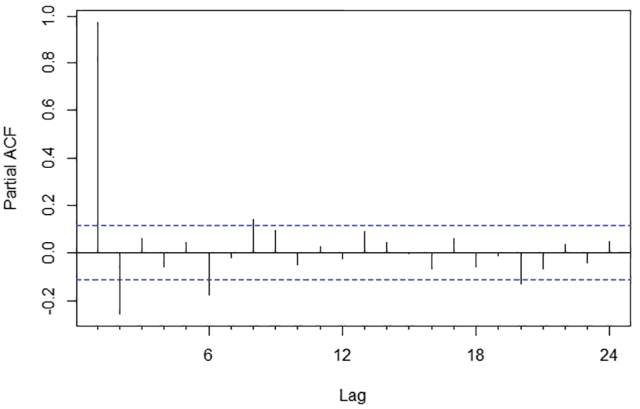


Fig. 5. PACF plot

2.2 Forecasting Models

In this paper five different forecasting models are used. All models were developed using *R* language and libraries. The classical ARIMA model was applied to this time series. Four different machine learning models will also be applied, namely Classification and Regression Tree, Random Forests, Support Vector Machines and Multivariate Adaptive Regression Splines. Considering the machine learning frameworks, different models were developed for each of the forecasting horizons. This means that a model will be trained to forecast prices one month in advance, another model to forecast two months in advance, and so on. As most of these models have parameters that have to be fixed *a priori*, different parameters were tested and the model that presented the best cross validation (CV) error in the training set was chosen. The CV error is calculated by considering, within the training set, a sample set that is not used for training but only for testing the model. CV error was also used for deciding how

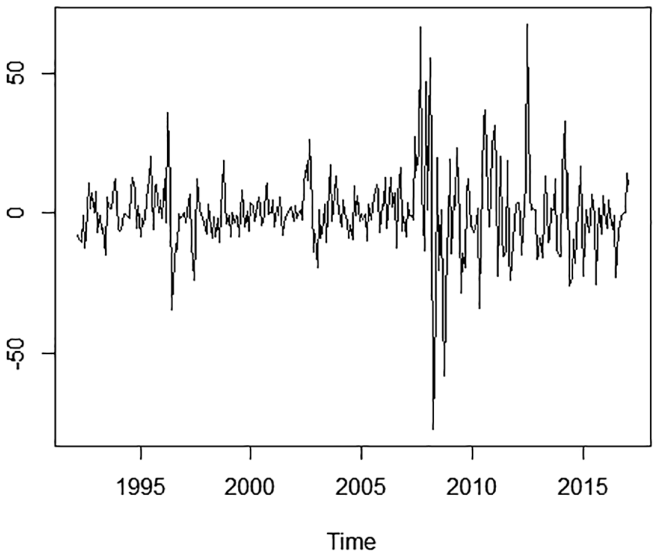


Fig. 6. First order differencing

many time lags should be considered in the input variables set. A short description of each of the models is presented next. It should be noticed that, unlike ARIMA models, machine learning models do not need to consider assumptions regarding the stationarity of the time series.

ARIMA

An ARIMA(p, d, q) is an autoregressive integrated moving average model where p represents the number of autoregressive terms, d represents the number of times the time series was differenced in order to become stationary and q is the number of moving average terms. Different models are obtained with different values of p, d and q. The model considered was the best one according to the Akaike Information Criteria (AIC), returned by using the function *auto.arima* from the *R* forecast package. Both seasonal and non-seasonal models were tested. The best model was a non-seasonal ARIMA(0, 1, 1). The obtained residuals are depicted in Fig. 7. Figures 8 and 9 present the ACF and PACF plots for the residuals. None of the autocorrelations and partial autocorrelations is individually statistically significant. The Box-Pierce presents a p-value equal to 0.79. The Ljung-Box statistics present a p-value equal to 0.67. High p-values allow us to conclude that the residuals estimated are purely random.

Classification and Regression Tree

CART models are represented by binary trees with different splitting rules in non-leaf nodes [19]. In each node, the algorithm will decide which variable gives the best split, by using a given criterion that corresponds to the minimization of a cost function (like the maximization of between-groups sum-of-squares error, for instance). The leaf nodes of the tree represent a given output variable that is used for making the prediction. After building the tree, it is usual to run a pruning algorithm, that will try to

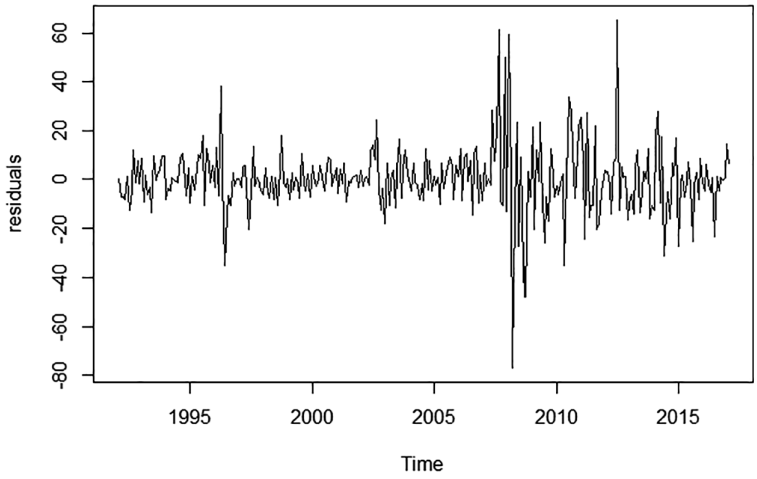


Fig. 7. ARIMA residuals



Fig. 8. In–sample and out-of-sample data

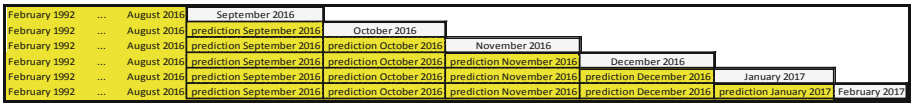


Fig. 9. Incorporating previous predicted values in longer term forecasts: grey cells represent forecasted values

remove leaf nodes if they do not contribute to a decrease in the cost function, aiming at achieving a less complex model.

To make a forecast considering a vector of new and unseen input variables, it is only necessary to go from the top to the bottom of the tree, respecting the rules in each node. There is one and only one path from the root of the tree to each one of the leaves. Package *rpart* has been used in the computational tests performed.

Random Forests

Random Forests can be interpreted as an ensemble of predictors. They are a combination of tree predictors such that each tree is constructed based on a sampling of the available training samples [20]. Each tree in the forest is also built considering a random selection of input variables that will determine the splitting rules in each node. Package *randomForest* for *R* has been used in the computational tests performed.

Support Vector Machines
Support Vector Machines are supervised learning models that were initially used as linear classifiers. They use kernels to extend data into a high dimensional feature space to improve the classification performance [21]. SVM can be applied to both classification problems and regression problems. SVM implementations require the user to define some parameters (the kernel function and a cost parameter that determines the trade off between model complexity and allowed deviations) [22]. Kernels usually considered are the linear kernel, the Gaussian basis function, the polynomial kernel, the Bessel function, the Laplace radial basis function. When there is no prior knowledge about the data, Gaussian, Laplace basis function and Bessel kernels are considered to be general-purpose kernels and thus an appropriate choice [23]. Package *Kernlab* was used in the computational tests performed.

Multivariate Adaptive Regression Splines
MARS has been first introduced by Friedman [24], and it is described by its author as a "flexible nonparametric regression modelling". The method produces continuous models with continuous derivatives, which is a differentiating aspect from the recursive partitioning approaches. MARS builds models that are weighted sums of basis functions, such that each basis function is either a constant, a function of the form max(0, x-constant) or max(0, constant-x), or a product of two or more of these functions. The algorithm begins by considering one single region. Then this region is recursively split, by defining a basis function, a predictor variable and a split point. The cost function to be minimized is the lack of fit of the model. Package *earth* has been used in the computational tests performed.

2.3 Methods

The forecasting models will be tested in out-of-sample data only. Prices from February 1992 to February 2012 will be used to estimate the models parameters. These models will then be used to forecast price values for the next six months (March 2012–August 2012). Then the in-sample time window will consider one more month (including March 2012). The models parameters are again estimated. New forecasts will be done for the next six months (April 2012–September 2012). The process will continue until the end of the available dataset. The last forecasting window will consider months from September 2016 to February 2017, being the models built with data until August 2016. This means that the models will be evaluated considering out-of-sample data, in a total of five years (60 months) predictions. Figure 8 illustrates the operation of the growing sliding window.

It is possible to calculate different forecasting metrics to assess the performance of the different models. Consider that y_t is the observed value at period t, $\hat{y}_t$ is the forecasted value for that time period, made k periods in advance, and that there are forecasts for periods $t \in T$. Mean relative absolute error (MAE) can be calculate as $\sum_{t \in T} \frac{|y_t - \hat{y}_t|}{y_t} / \#T$. Maximum relative absolute error (MaxAE) can be calculated as $max\left\{\frac{|y_t - \hat{y}_t|}{y_t}, \forall t \in T\right\}$. Root Mean Square Error (RMSE) is calculated as $\sqrt{\frac{\sum_{t \in T}(y_t - \hat{y}_t)^2}{\#T}}$. These error metrics are calculated for all the forecasting futures considered (one to six

months). The ability of correctly forecasting whether the price will increase or decrease in the future is also going to be considered.

3 Computational Results

Each of the models developed was tested in out-of-sample data, as explained in Sect. 2.3. The capability of predicting wheat prices one up to six months ahead is evaluated using different performance measures. All models were also compared with a naïve model that considers the prediction $\hat{y}_t = y_{t-k}$, meaning that the prediction made k months ahead will simply be equal to the current price.

Table 1 presents the results considering MAE. Table 2 presents MaxAE results, and Table 3 RMSE results. The best value found is highlighted in each one of the tables.

Table 1. Mean Absolute Error (MAE)

	months ahead					
	1	**2**	**3**	**4**	**5**	**6**
ARIMA	4,34%	9,49%	12,87%	16,05%	18,73%	21,29%
CART	5,43%	9,06%	13,54%	16,17%	16,86%	18,95%
RF	4,35%	9,15%	12,55%	15,25%	19,31%	21,74%
SVM	2,55%	5,40%	8,07%	11,72%	14,35%	15,88%
MARS	2,22%	5,67%	8,09%	11,83%	15,48%	18,60%
Näive	4,54%	7,52%	9,68%	11,78%	13,12%	14,55%

As expected, the models performance deteriorates as the forecasting horizon increases. No model is the "best" model under all the possible performance criteria. We can, however, conclude that SVM and MARS models present the best results for one to four months forecasts. It is interesting to note that the näive model achieves the best results for five and six months forecasts.

Models were also tested for the capacity of accurately forecasting if the future time series value would be greater or lower than the current one. Accuracy results are presented in Table 4. In this case, SVM presents the best results for all forecasting horizons. Results are very good in predicting the increase or decrease in future price compared with the current one for one and two months forecasts.

As it was possible to obtain very good results with MARS for one month forecasts, it makes sense to try to develop two to six months forecast models incorporating these predictions. This means that only one month prediction models are developed, since longer term predictions will consider as input variables the values already predicted for the shorter term forecasts. Figure 9 illustrates this situation. After having forecasts for the next month, it is possible to include these values as input data in longer term forecasts, so that two months forecasts still use the one month forecast model, and so on.

Table 2. Maximum Absolute Error (MaxAE)

	months ahead					
	1	**2**	**3**	**4**	**5**	**6**
ARIMA	21,58%	27,85%	37,66%	35,29%	45,21%	50,73%
CART	16,61%	23,55%	28,99%	41,52%	57,48%	61,70%
RF	15,79%	25,40%	27,22%	31,05%	46,23%	71,73%
SVM	8,89%	14,37%	24,76%	30,71%	35,04%	36,33%
MARS	6,77%	16,77%	23,10%	32,09%	39,05%	38,79%
Näive	21,90%	25,37%	27,15%	36,32%	35,28%	33,33%

Table 3. Root Mean Square Error

	months ahead					
	1	**2**	**3**	**4**	**5**	**6**
ARIMA	14,45	26,81	34,95	39,15	44,29	48,16
CART	15,57	21,97	32,71	39,10	41,05	43,92
RF	12,83	25,91	32,44	36,81	45,40	48,92
SVM	7,58	14,66	23,14	31,89	37,82	38,90
MARS	6,17	15,50	22,43	32,04	39,23	43,61
Näive	14,79	23,27	29,49	33,87	35,76	36,87

Table 4. Accuracy in Trend

	months ahead					
	1	**2**	**3**	**4**	**5**	**6**
ARIMA	55,93%	54,39%	50,91%	45,28%	50,98%	55,10%
CART	61,02%	50,88%	52,73%	49,06%	47,06%	48,98%
RF	69,49%	49,12%	58,18%	50,94%	47,06%	42,86%
SVM	86,44%	75,44%	69,09%	56,60%	58,82%	61,22%
MARS	84,75%	75,44%	65,45%	56,60%	47,06%	40,82%

Tables 5, 6, 7 and 8 show the performance metrics for this new approach, comparing with the initial MARS models and also with the best performance obtained earlier. It is possible to conclude that the performance for two to four months forecasts improves, since this methodology obtains the best results for MAE and RMSE.

Table 5. Mean Absolute Error (MAE)

	2	3	4	5	6
Best previous result	5,40%	8,07%	11,72%	13,12%	14,55%
MARS	5,67%	8,09%	11,83%	15,48%	18,60%
MARS including previous forecasted values	3,52%	7,58%	10,79%	13,24%	14,55%

Table 6. Maximum relative absolute error (MaxAE)

	2	3	4	5	6
Best previous result	14,37%	23,10%	30,71%	35,04%	33,33%
MARS	16,77%	23,10%	32,09%	39,05%	38,79%
MARS including previous forecasted values	14,04%	24,06%	35,52%	37,76%	47,27%

Table 7. Root Mean Square Error (RMSE)

	2	3	4	5	6
Best previous result	14,66	22,43	31,89	35,76	36,87
MARS	15,50	22,43	32,04	39,23	43,61
MARS including previous forecasted values	10,69	21,53	30,00	35,49	37,99

Table 8. Accuracy in Trend

	2	3	4	5	6
Best previous result	75,44%	69,09%	56,60%	58,82%	61,22%
MARS	84,75%	75,44%	65,45%	56,60%	47,06%
MARS including previous forecasted values	89,47%	74,55%	60,38%	52,94%	51,02%

Figures 10a to f illustrate the forecasting values obtained by these models for $t + k$ time period, with t the current time period and k the future time horizon, $k = 1, \ldots, 6$.

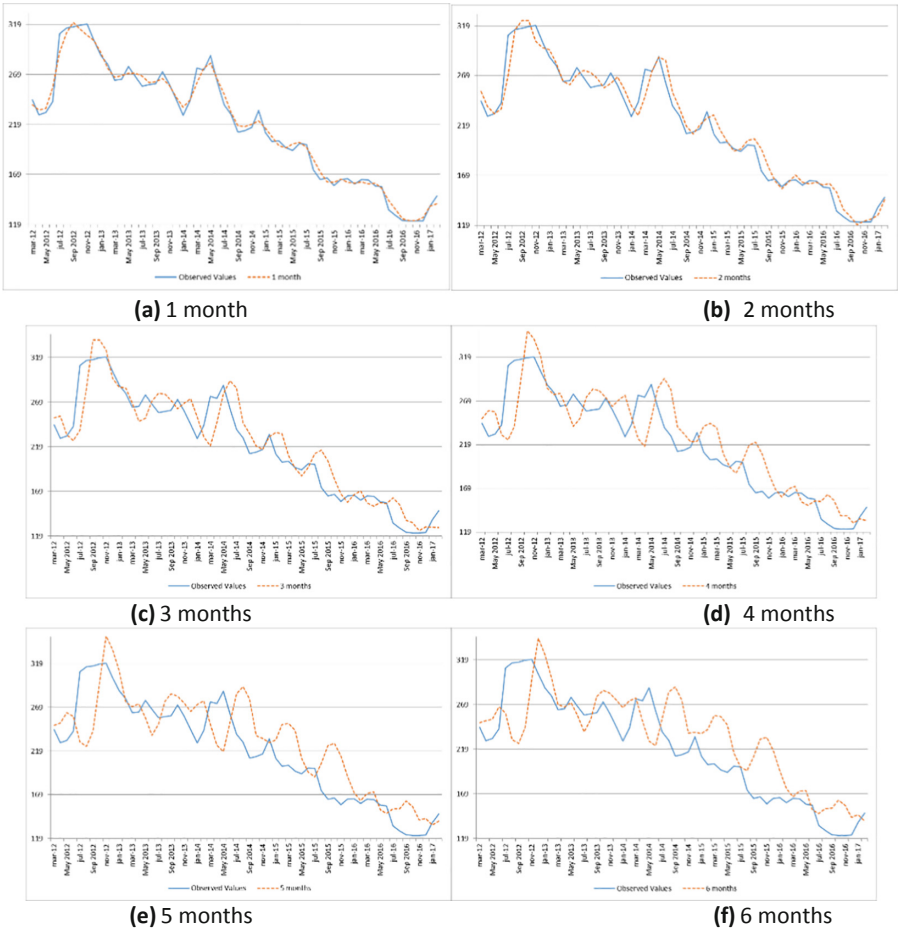


Fig. 10. Prediction based on one month forecast MARS models, incorporating previous predictions

Figure 11 illustrates the results obtained but from a different point of view: at time t it is possible to predict prices for $t + k$, for $k = 1, \ldots, 6$. Figure 11 considers the prediction of prices between September 2016 and February 2016, using data available until August 2016. As can be seen in the figure, the forecasting value is capturing very well the trend of the price evolution.

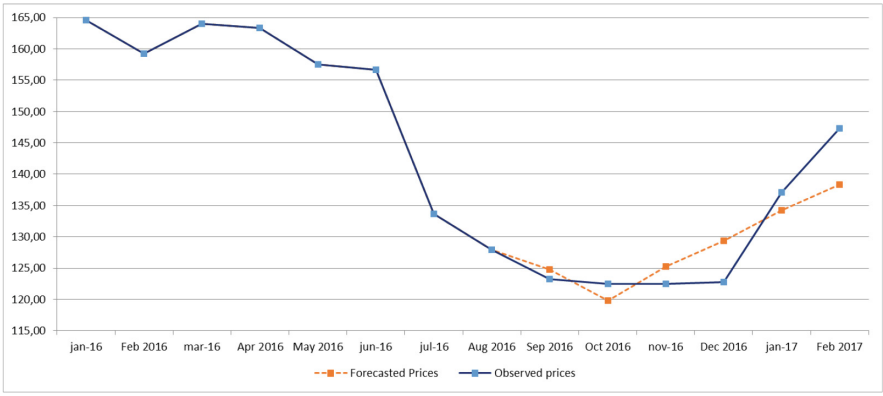


Fig. 11. Forecasts for the last six months of the available dataset

4 Conclusions

In this paper we have presented five different types of models to forecast wheat price one up to six months ahead, considering as predictor variables past time series values only. As far as the authors know, it is the first time that the investigated machine learning techniques are applied to this problem. The methodology used is different from other published works, since we aim at forecasting at the present time the prices for the next six months. All models were tested considering different performance criteria and using out-of-sample data.

Although it is not possible to select one approach as the best one under all criteria, the approach based on MARS that includes previous forecasted values seems to be the one with most consistent results.

As future work, it is possible to devise a forecasting model that considers an ensemble of different predictors. An ensemble can be understood as a collection of learning algorithms that are simultaneously used for making more reliable and accurate predictions than its individual components [25]. This collection can consider different machine learning based models, or the same type of models but trained using different sample sets or different input variables. It is also possible to consider the application of classification algorithms capable of accurately predicting the future increase or decrease in the price.

Acknowledgements. This work has been supported by the Fundação para a Ciência e a Tecnologia (FCT) under project grant UID/Multi/00308/2019.

References

1. Curtis, B.C., Rajaram, S., Gómez, M.: Bread wheat: improvement and production: Food and Agriculture Organization of the United Nations (FAO) (2002)
2. Yang, J., Zhang, J., Leatham, D.J.: Price and volatility transmission in international wheat futures markets. Ann. Econ. Finan. **4**, 37–50 (2003)
3. Brunner, A.D.: El Nino and world primary commodity prices: warm water or hot air? Rev. Econ. Stat. **84**, 176–183 (2002)
4. Hill, H.S.J., et al.: Implications of seasonal climate forecasts on world wheat trade: a stochastic, dynamic analysis. Can. J. Agr. Econ./Revue canadienne d'agroeconomie **52**, 289–312 (2004)
5. Ubilava, D.: The ENSO effect and asymmetries in wheat price dynamics. World Dev. **96**, 490–502 (2017)
6. Ramirez, O.A., Fadiga, M.: Forecasting agricultural commodity prices with asymmetric-error GARCH models. J. Agr. Resource Econ. **28**(1), 71–85 (2003)
7. Benavides, G.: Price volatility forecasts for agricultural commodities: an application of historical volatility models, option implieds and composite approaches for futures prices of corn and wheat. SSRN (2004)
8. Bessler, D.A., Yang, J., Wongcharupan, M.: Price dynamics in the international wheat market: modeling with error correction and directed acyclic graphs. J. Reg. Sci. **43**, 1–33 (2003)
9. Roche, M.J., McQuinn, K.: Grain price volatility in a small open economy. Eur. Rev. Agr. Econ. **30**, 77–98 (2003)
10. Hoffman, L., Irwin, S.H., Toasa, J.: Forecast performance of futures price models for corn, soybeans, and wheat. Agricultural & Applied Economics Association, Milwaukee, WI (2007)
11. Jumah, A., Kunst, R.M.: Seasonal prediction of European cereal prices: good forecasts using bad models? J. Forecast. **27**, 391–406 (2008)
12. Arshad, F.M., Hameed, A.A.A.: The long run relationship between petroleum and cereals prices. Glob. Econ. Finan. J. **2**, 91–100 (2009)
13. Algieri, B.: A roller coaster ride: an empirical investigation of the main drivers of the international wheat price. Agr. Econ. **45**, 459–475 (2014)
14. Chen, Y.-C., Rogoff, K.S., Rossi, B.: Can exchange rates forecast commodity prices? Quart. J. Econ. **125**, 1145–1194 (2010)
15. Adjemian, M.K., Janzen, J., Carter, C.A., Smith, A.: Deconstructing Wheat Price Spikes: A Model of Supply and Demand, Financial Speculation, and Commodity Price Comovement. United States Department of Agriculture, Economic Research Service (2014)
16. Ahumada, H., Cornejo, M.: Forecasting food prices: the case of corn, soybeans and wheat. Int. J. Forecast. **32**, 838–848 (2016)
17. Khamis, A., Abdullah, S.: Forecasting wheat price using Backpropagation and NARX Neural Network. Int. J. Eng. Sci. **3**, 19–26 (2014)
18. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.: STL: a seasonal-trend decomposition procedure based on loess. J. Official Stat. **6**, 3–73 (1990)
19. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees. CRC Press, Boca Raton (1983)
20. Breiman, L.E.O.: Random forests. Mach. Learn. **45**, 5–32 (2001)
21. Li, D.-C., Liu, C.-W.: A class possibility based kernel to increase classification accuracy for small data sets using support vector machines. Expert Syst. Appl. **37**, 3104–3110 (2010)

22. Chapelle, O., Vapnik, V.: Model selection for support vector machines. In: Solla, S., Leen, T.K., Miller, K.-R. (eds.) Advances in Neural Information Processing Systems 12, pp. 230–236. MIP Press (2000)
23. Hornik, K., Meyer, D., Karatzoglou, A.: Support vector machines in R. J. Stat. Softw. **15**, 1–28 (2006)
24. Friedman, J.H.: Multivariate adaptive regression splines (with discussion). Ann. Stat. **19**, 79–141 (1991)
25. Re, M., Valentini, G.: Ensemble methods: a review. In: Srivastava, A.N., Scargle, J.D., Ali, K., Way, M.J. (eds) Data Mining and Machine Learning for Astronomical Applications, pp. 1–40. Chapman & Hall (2011)