Forecasting Agricultural Commodity Prices using Hybrid Neural Networks

Tamer Shahwan[1] Martin Odening[2]

[1]Humboldt-Universität zu Berlin, School of Business and Economics, Institute of Banking, Stock Exchanges and Insurance, D-10178 Berlin, Germany and Faculty of Agriculture and Horticulture, Department of Agricultural Economics and Social Sciences D-10099 Berlin, Germany
[2]Humboldt-Universität zu Berlin, Faculty of Agriculture and Horticulture, Department of Agricultural Economics and Social Sciences, D-10099 Berlin, Germany

Abstract

Traditionally, the autoregressive integrated moving average (ARIMA) model has been one of the most widely used linear models in time series forecasting. However, the ARIMA model cannot easily capture non-linear patterns. In the last two decades artificial neural networks (ANN) have been proposed as an alternative to traditional linear models, particularly in the presence of nonlinear data patterns. Recent research suggests that a hybrid approach combining both ARIMA models and ANN can lead to further improvement in the forecasting accuracy compared with pure models. In this paper, a hybrid model that combines a seasonal ARIMA model and an Elman neural network (ENN) is used to forecast agricultural commodity prices. A genetic algorithm (GA) is employed to determine the optimal architecture of the ANN. It turns out that the out-of-sample prediction can be improved slightly with the hybrid model.

Keywords: Time Series Forecasting; ARIMA; Elman Neural Network; Genetic Algorithms; Hybrid System; Commodity Prices.

1. Introduction

Time series analysis is an important approach to forecasting in which past observations of the same variable are collected and analyzed to develop a model describing the underlying data generating process. Much effort has been devoted over the past decades to the development of time series forecasting models. One of the most important and widely used time series models are autoregressive integrated moving average (ARIMA) models, which serve as a benchmark model for creating linear models because of their theoretical elaborateness and accuracy in short term forecasting [7].

Recently, artificial neural networks (ANN) have been proposed as an alternative forecasting technique. Zhang et al. [24] summarize the different applications of neural networks for forecasting.

There is a number of studies in which neural networks are used to address financial economic problems. For instance, Kaastra and Boyd [8] provide a general introduction of how to model financial and economic time series using ANN. Yao et al. [20] use backpropagation neural networks for forecasting the option prices of Nikkei 225 index futures. For predicting agricultural commodity prices, there are few articles using artificial neural networks. For instance, Kohzadi et al. [9] use backpropagation networks to model monthly live cattle and wheat prices and compare the results with that obtained by ARIMA model. They conclude that neural networks perform better in terms of mean absolute error (MAE). Moreover, neural networks are also better in capturing turning points. Richards et al. [16] compare two methods of estimating a reduced form model of fresh tomato marketing margins: an econometric and an artificial neural network approach. The neural network is able to forecast with approximately half the mean square error of the econometric model, but both are equally adept at predicting turning points in the time series.

In this paper, we will investigate the ability of a hybrid approach combining ARIMA and evolutionary Elman neural network to time series forecasting against each stand-alone model. The motivation of this approach is largely due to the fact that a real-world problem is often complex in nature and any individual model may not be able to capture different patterns equally well [23].

Using hybrid models could complement each other in capturing patterns of data sets and could improve the forecasting accuracy. The literature on hybrid models has expanded rapidly since the early work of Bates and Granger [1]. Clemen [4] provides a comprehensive review and annotated bibliography in this area. Wedding and Cios [18] described a combining methodology using radial basis function networks and the Box-Jenkins models. Menezes et al. [11] offer a good guideline for combined forecasting. They concluded that the use of different criteria is common in evaluating the forecasting performance. The properties of individual forecasts errors can strongly influence the characteristics of the combination's error. Pendharkar [14] proposes a hybrid evolutionary neural approach for binary classification. Zhang [23] combines the ARIMA and feedforward neural networks models in forecasting. Pai and Lin [13] combine ARIMA and support vector machines model in stock price forecasting. The current study will examine the forecasting accuracy of a hybrid model of ARIMA and ENN to forecast agricultural commodity prices.

2. A Hybrid forecast model and its components

2.1. Box-Jenkins model

The familiar Box-Jenkins approach combines two types of processes: autoregressive (AR) and moving average (MA). The general class of ARMA (*p,q*) model has the following form [17]:

$$(1-\phi_1 \mathrm{B} - .. - \phi_p \mathrm{B}^p) y_t = \phi_0 + \left(1-\theta_1 \mathrm{B} - .. - \theta_q \mathrm{B}^q\right) a_t \quad (1)$$

where $\{a_t\}$ is a white noise series, p and q are non negative integers, B is the back shift operator with $\mathrm{B}(y_t) = y_{t-1}$ where y_t denotes the monthly prices. This model to be meaningful, we need $\phi_1 \neq \theta_1$, otherwise there is a cancellation in the equation and the process reduces to a white noise series. Since many time series are nonstationary, differencing one or more times is required. This leads to an ARIMA (*p,d,q*) model [15]:

$$(1-\phi_1\mathrm{B}-\phi_2\mathrm{B}^2-....-\phi_p\mathrm{B}^p)(1-\mathrm{B})^d y_t = \left(1-\theta_1\mathrm{B}-\theta_2\mathrm{B}^2-....-\theta_q\mathrm{B}^q\right)a_t \quad (2)$$

Furthermore, the regular and seasonal components of a time series can be captured by a general multiplicative ARIMA model [12]:

$$\phi_p(\mathrm{B})\Phi_P(\mathrm{B}^s)(1-\mathrm{B})^d\left(1-\mathrm{B}^s\right)^D y_t = \theta_q(\mathrm{B})\Theta_Q(\mathrm{B}^S)a_t \quad (3)$$

where B^s is the seasonal back shift operator. $\Phi_p(\mathrm{B}^s) = (1-\phi_{1s}\mathrm{B}^s - \phi_{2s}\mathrm{B}^{2s} - ......- \phi_{ps}\mathrm{B}^{ps})$ is the seasonal auto-regressive process. $\Theta_Q(\mathrm{B}^s) = (1-\Theta_s\mathrm{B}^s - \Theta_{2s}\mathrm{B}^{2s} - ....- \Theta_{Qs}\mathrm{B}^{Qs})$ is the seasonal moving average process. In general, s equals 4 or 12. P, D, Q have values of 0, 1 or 2. A useful notation to describe the orders of the various components in this multiplicative model is given by $(p,d,q)\times(P,D,Q)^s$ where P is the seasonal level of auto-regressions, D is the seasonal level of differences and Q is the seasonal level of moving average.

Basically, the application of ARIMA models consists of three phases: model identification, parameter estimation and diagnostic checking. The identification step requires an intensive data analysis where expert judgment must be exercised to interpret the behaviour of the autocorrelation function (ACF) and the partial autocorrelation function (PACF). However, any significant nonlinearity limits the application of ARIMA models. Therefore, a hybrid model using also an Elman neural network is proposed to deal with a non-linear pattern possibly present in the data.

2.2. Elman Neural Networks

The Elman neural network (ENN) is one kind of globally feedforward locally recurrent network model proposed by Elman [5]. ENN can be considered as an extension of Multilayer perceptron (MLP) with an additional input layer (state layer) that receives a feedback copy of the activations from the hidden layer at the previous time step. These context units in the state layer in the Elman network make it sensitive to the history of input data, which is essentially useful in dynamical system modelling [19].

Fig. 1 depicts the idea of ENN where the activations in the hidden layer at time *t*-1 are copied into the context vector as input to the network at time *t* [3].

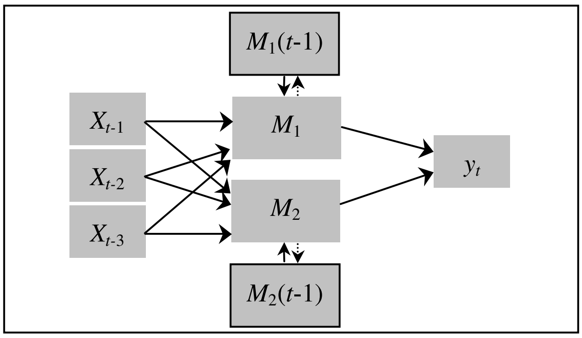

Fig. 1: Elman Neural Network from McNelis (2005)

An ENN with a "tanh sigmoidal activation function" has the following structure [10]:

$$n_{k,t} = w_{k,o} + \sum_{i=1}^{i*} w_{k,i} x_{i,t} \quad (4)$$

$$M_{k,t} = \frac{e^{n_{k,t}} - e^{-n_{k,t}}}{e^{n_{k,t}} + e^{-n_{k,t}}} + \sum_{k=1}^{k*} \phi_k M_{k,t-1} \quad (5)$$

$$y_t = \gamma_0 + \sum_{k=1}^{k*} \gamma_k M_{k,t} \quad (6)$$

where x_i are input variables $(1,2,.......,i*)$. $k*$ is the number of neurons, $w_{k,0}$ is a constant term, $w_{k,i}$ are the synaptic weights of input variables, $n_{k,t}$ is a linear combination of these input variables observed at time t, γ_0 is the constant term, γ_k is the coefficient scalar between the hidden and the output layer and ϕ_k are weights between context units and hidden units. Hence, the output (y_t) depends not only on the new inputs but also on the preceding context units.

2.3. Hybrid models

The behaviour of commodity prices may not easily be captured by stand-alone models because time series could include a variety of characteristics such as seasonality, heteroskedasticity or non-gaussian error. It is also difficult for a forecaster to determine whether the time series under study is generated from a linear or a non-linear underlying process [23]. Therefore, a hybrid model having both linear and non-linear modelling abilities could be a good alternative to forecast commodity prices. Following Zhang [23] and Yu et al. [22], we assume that S_t is a time series that is composed of a linear part L_t and a non-linear part N_t:

$$S_t = L_t + N_t \quad (7)$$

Both L_t and N_t are estimated from the data by a two-step-procedure: First, we use an ARIMA model to capture the linear component. The residuals e_t from this linear model defined by

$$e_t = S_t - \hat{L}_t \quad (8)$$

can be considered as an approximation of the nonlinear part N_t, where $\hat{L}_t$ is the estimated value of the ARIMA model at time *t*. In a second step the residuals are modeled by an evolutionary Elman neural network, the nonlinear pattern in the residuals can be discovered. With *m* input nodes, these residuals are modeled as follows

$$e_t = f(e_{t-1}, e_{t-2}, e_{t-3}, \ldots, e_{t-m}) + \varepsilon_t \quad (9)$$

where f is a nonlinear function determined by the neural network and ε_t is the random error. Finally, the combined forecast is

$$\hat{S}_t = \hat{L}_t + \hat{e}_t \quad (10)$$

with $\hat{e}_t$ being the forecast of the residual using the ENN.

3. Empirical analysis

3.1. Data

Two data sets, monthly hog prices and canola prices from Germany, are now used to investigate the effectiveness of the proposed hybrid model compared to an ARIMA models and an ENN[1]. We assess the forecasting performance by means of an out-of-sample-technique. Each time series is divided into a training set and testing set. The training set is used for model specification and then the test set is used to evaluate the established model. In order to achieve a good generalization of the ENN, a cross-validation approach is adopted. That means the training set is further portioned into two subsets: 80 percent of the training set is assigned to the estimation and the remaining 20 percent is assigned to the validation subset [6]. Before estimation the data are normalized to the range between [-1,1]. The decomposition of the two data sets is given in table 1.

The measurement of prediction performance of the three models is based on one-step-ahead-forecasts over the test set. The mean absolute percentage error (MAPE) and Theil's are employed to measure the forecasting errors:

$$MAPE = \frac{100}{T} \sum \left| \frac{\hat{y}_t - y_t}{y_t} \right| \quad (11)$$

$$Theil's\,U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T} (\hat{y}_t - y_t)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} (\hat{y}_t)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T} (y_t)^2}} \quad (12)$$

where y_t and $\hat{y}_t$ are the actual and the predicted price, respectively, at time *t* and *T* is the number observations in the test set.

Series	Total Sample	Training set	Cross-validation set	Test set
Hog prices (€/kg slaughter weight)	1972-2003 (384)	1972-1999 (336)	20% of training set (67)	2000-2003 (48)
Canola prices[2] (€/ton)	1992-july 2004 (151)	1992-2001 (120)	20% of training set (24)	2002-july 2004 (31)

Table 1: Sample decomposition

3.2. Specification of the forecasting models

The estimated ARIMA model for the hog data has the structure $(2,0,0)H(1,0,1)_{12}$. The canola data is fitted best with an autoregressive model (AR) of order 3. Next, we test for the presence of nonlinearieties in the data. If nonlinearities are statistically significant, then choosing a class of nonlinear models like ANN might improve the predictive power. In this context, we apply the Ljung-Box-Q-statistic to the squared residuals of the two ARIMA models.

Time series	Q (5)	Q (10)	Q (15)	Q (20)	Q (24)
Hog Prices	36.27	43.97*	51.30*	81.18*	89.53*
Canola Prices	6.10	6.11	6.12	6.12	6.13

* Denotes statistical significance at the 5% level

Table 2: The Ljung-Box statistics for the squared residuals of an ARMA model

The results in table 2 show that there is evidence for the presence of GARCH effects (i.e. conditional heteroscedasticity) in the hog prices. No such effect occurs in the canola data. Conditional heteroskedasticity implies that the underlying times series is nonlinear in variance. Baillie and Bollerslev [2] show that conditional heteroscedasticity changes the mean square error (MSE) of the predictor. Yim [21] successfully applies ANN to data that show conditional het-

[1] The data are obtained from the "Zentrale Markt- und Preisberichtstelle GmbH- Marktberichtstelle" (ZMP), Berlin.

[2] Monthly canola prices include missing data in the following months: July 92, June 93, April 94, Mai 94, June 94, July 94, July 95, July 96, July 97, July 98, July 2000, July 2002 and July 2003. The interpolation method is used for replacing the missing observations.

eroscedasticity. In his study ANN prove to be superior to ARIMA-GARCH models. Due to this finding we conjecture that ENN will outperform the traditional ARIMA models particularly for the hog price data.

The implementation of the ENNs requires to specify a large number of parameters. In order to specify the ENNs, we apply a mixture of different methods. First, the number of neurons, m, in the input layer is heuristically determined by the number of autoregressive (AR) terms of the ARIMA models. Hence, the variables (y_{t-1}, y_{t-2}, y_{t-12}, y_{t-13}, y_{t-14}) are used as inputs to the ENN in the case of the hog prices. The corresponding input variables for the canola prices are (y_{t-1}, y_{t-2}, y_{t-3}). The hyperbolic tangent function is chosen as a transfer function between the input and the hidden layer. The identity transfer function connects the hidden and the output layer. Second, the weight decay method as pruning technique is used to reduce the weights and to find a parsimonious structure for the ENN. Finally, a genetic algorithm (GA) as a global search procedure supports configuration process of the ENN. In particular, the number of hidden nodes, the value of learning rate, the momentum term and the weight decay constant are genetically optimised. The ENN is trained by backpropagation. Batch updating is chosen as the sequence, in which the patterns are presented to the network. As mentioned above, the cross-validation approach is used to determine the optimal number of training epochs. That means the training procedure is terminated as soon as the mean square error of the cross-validation set increases, since this indicates that the network has begun to overtrain. This prevents the ENNs from memorizing unnecessary noise in time series.

To set up the hybrid models, the linear part of the price data is filtered by the aforementioned AIRMA models. The residuals of the ARIMA models are then fed into an ENN as explained in section 2.3. The ENNs in the hybrid models have the same structure as the ENNs that are directly used for analyzing the price data.

3.3. Results

ARIMA models are estimated using Statgraphics Plus software. The software package NeuroSolutions V4.32 (NeuroDimensions Inc., Gainesville, FL) was employed for the estimation of the ENNs. NeuroSolutions also includes a GA to optimize the structure of ANN.

The out-of-sample forecast performance of the ARIMA model, the ENN and the hybrid model for both hog and canola price data are reported in table 3 (see also figures A1 and A2 in the appendix).

The figures in table 3 indicate that only slight differences in the predictive power of the various methods exist. With regard to the canola price data, no improvement of the forecast accuracy can be achieved by using an ENN instead of a traditional ARIMA model. The latter is even better than the former. This raises some doubt, if the optimal specification of the ENN has been identified for this time series. The joint use of both methods in the hybrid model leads to a minor increase of the accuracy compared with the single ARIMA model. In accordance with our initial guess the ENN outperforms the ARIMA model for the hog prices, which obey a more complex and nonlinear pattern. The benefits from merging the two methods are not clear. While the MAPE of the hybrid model is smaller than for the ENN the opposite is true for Theil's U. It should be mentioned that the differences in the prediction errors are not statistically significant at the 5% level based on a t-test.

Time series	MAPE (Rank)	Theil's U (Rank)
ARIMA		
Hog prices	4.5463 (3)	0.030431 (3)
canola prices	3.0841 (2)	0.022246 (2)
ENN		
Hog prices	4.4741 (2)	**0.029691** (1)
Canola prices	3.2333 (3)	0.022469 (3)
Hybrid Model		
Hog prices	**4.4674** (1)	0.03003 (2)
Canola prices	**3.0531** (1)	**0.021896** (1)

Bold letters indicate minimal errors

Table 3: Comparison of forecast errors

4. Conclusions

In this study we explore the usefulness of ANN for short term forecasting of agricultural commodity prices. Traditional ARIMA models serve as a benchmark for the evalution. Moreover, we investigate whether a hybrid model is beneficial compared to the application of single forecast methods as the recent literature suggests. Our results are not unambiguous. Obviously, the potential gain of ANN and hybrid models seem to depend on the characteristics of the time series under consideration. A more complex time series justifies the use of an ANN. However, the greater flexibility of this model class and its ability to handle nonlinear data patterns comes at the cost of a more demanding specification procedure. While the Box-Jenkins approach is straightforward, there is no simple clear-cut method to determine the optimal structure of the ANN. Zhang et al. [24] concede that the design of an ANN is more an art than a science. That means the danger of missspecifying an ANN is higher than for an ARIMA model. This may erode the potential superiority of the ANN and it also hampers the ability to generalize during comparisons between ANN and other forecast methods.

Furthermore, our calculations confirm that a hybrid model *may* reduce the prediction errors of single forecasting methods. But the question remains, when it is actually worthwhile to apply such a complex model, whose set-up-costs are relatively high. In order to identify situations more clearly where the use of hybrid models is recommended, comprehensive and systematic experiments with simulated data are necessary. We suggest such simulations experiments for further research.

5. References

[1] J.M. Bates and C.W.J. Granger, "The Combination of Forecasts", *Operations Research 20*, pp. 451-468, 1969.

[2] R. T. Baillie and T. Bollerslev, "Prediction in Dynamic Models with Time-dependent Conditional Variance", *Journal of Econometrics 52*, pp. 91-113, 1992.

[3] V. Cho, "A Comparison of Three Different Approaches to Tourist Arrival Forecasting", *Tourism Management 24*, pp. 323-330, 2003.

[4] R. Clemen, "Combining forecasts: A Review and Annotated Bibliography with Discussion", *International Journal of Forecasting 5*, pp. 559-608, 1989.

[5] J.L. Elman, "Finding Structure in Time", *Cognitive Science 14*, pp. 179-211, 1990.

[6] S. Haykin, "Neural Networks: A Comprehensive Foundation", Second Edition, Prentice-Hall, Inc., New Jersey, 1999.

[7] W.C. Jhee and M.J. Shaw, "Time Series Prediction using Minimally Structured Neural Networks: An Empirical Test", In: R.R. Trippi and E. Turban, "*Neural Networks in Finance and Investing*", Revised Edition, Irwin, Chicago, pp. 763-776, 1996.

[8] I. Kaastra and M. Boyd, "Designing a Neural Network for Forecasting Financial Economic Time Series", *Neurocomputing 10*, pp. 215-236, 1996.

[9] N. Kohzadi, M.S. Boyd, and B. Kermanshahi, "A Comparison of Artificial Neural Network and Time Series Models for Forecasting Commodity Prices", *Neurocomputing 10*, pp. 169-181, 1996.

[10] P. D. McNelis, "Neural Networks in Finance: Gaining predictive Edge in the Market", Elsevier Academic press, 2005.

[11] L.M. Menezes, D.W. Bunn, and J.W. Taylor, "Review of Guidelines for the Use of Combined Forecasts", *European Journal of Operational Research* 120, pp. 190-204, 2000.

[12] C.R. Nelson, "Applied Time Series Analysis for Managerial Forecasting", Holland-Day, Inc., San Francisco, 1973.

[13] P.-F. Pai, C.-S. Lin, "A Hybrid ARIMA and Support Vector Machines Model in Stock Price Forecasting", *Omega*, 2004, article in press.

[14] P.C. Pendharkar, "An Empirical Study of Designing and Testing of Hybrid Evolutionary-Neural Approach for Classification", *Omega 29*, pp. 361-374, 2001.

[15] R. S. Pindyck and D.L. Rubinfeld, "Econometric Models and Economic Forecasts", McGraw-Hill, Inc, 1991.

[16] T.J. Richards, P.M. Patterson, and P.V. Ispelen, "Modelling Fresh Tomato Marketing Margins: Econometrics and Neural Networks", *Agricultural and Resource Economics Review 27*, pp. 186-199, 1998.

[17] R.S. Tsay, "Analysis of Financial Time Series", John Wiley & Sons, Inc, 2002.

[18] D.K. Wedding and K.J. Cios, "Time Series Forecasting by Combining RBF networks, Certainty Factors, and the Box-Jenkins Models", *Neurocomputing 10*, pp. 149-168, 1996.

[19] F. Yang, H. Sun, Y. Tao, and B. Ran, "Temporal Difference Learning with Recurrent Neural Network in Multi-step ahead Freeway Speed Prediction", TRB 2004 Annual Meeting, 2004.

[20] J. Yao, Y. Li, and C.L. Tan, "Option Prices Forecasting Using Neural Networks", *Omega 28*, pp. 455-466, 2000.

[21] J. Yim, "A Comparison of Neural Networks with Time Series Models for Forecasting Returns on a Stock Market Index", Proc.Of the 15th International Conference on Industrial and Engineering. Applications of Artificial Intelligence and Expert Systems (IEA/AIE), Cairns, Australia, 2002.

[22] L. Yu, S. Wang, and K.K. Lai, "A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR and ANN For Foreign Exchange Rates", *Computers & Operations Research*, 2004, article in press.

[23] G.P. Zhang, "Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model", *Neurocomputing 50*, pp. 159-175, 2003.

[24] G. Zhang, B.E. Patuwo, and M.Y. Hu, "Forecasting with Artificial Neural Networks: The State of The Art", *International Journal of Forecasting 14*, pp. 35-62, 1998.

Contact person:

Tamer Shahwan
phone: +49-30-2093 6435; fax: +49-30-2093 6465
E-mail: tamer.shahwan@agrar.hu-berlin.de

Appendix

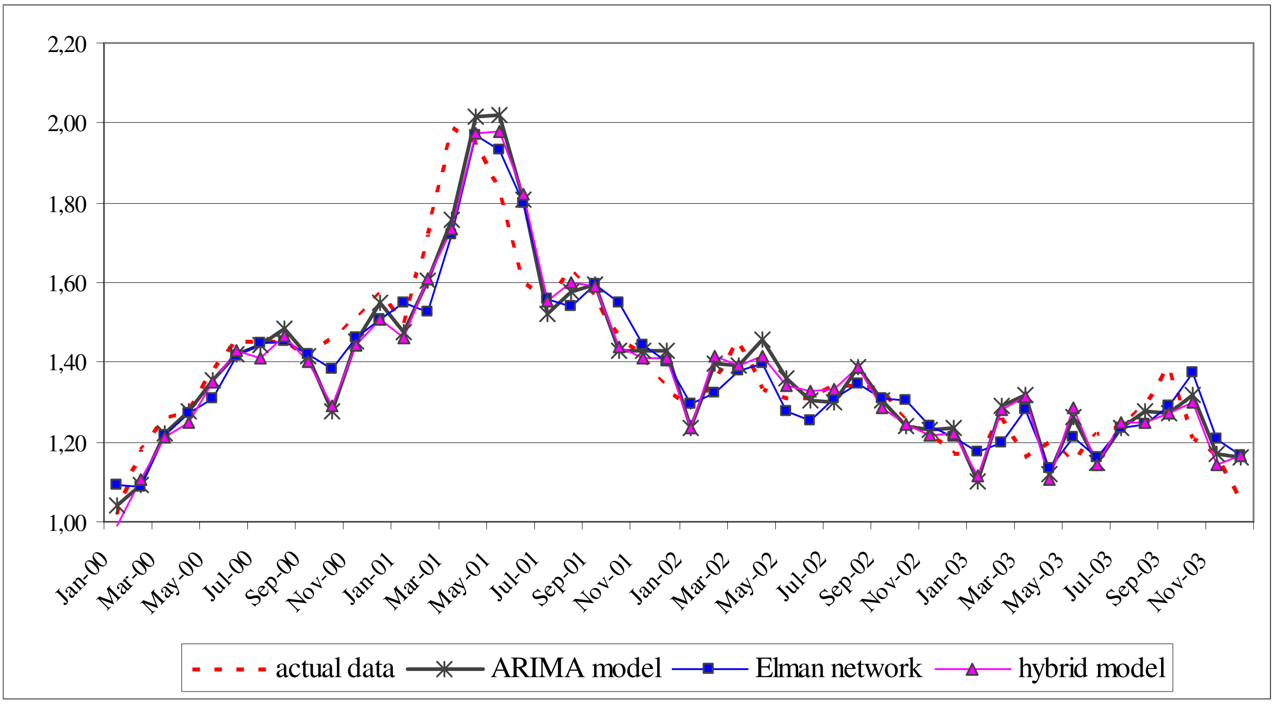

Fig. A 1: One-step-ahead forecasts of hog prices (Germany, January 2000 to December 2003)

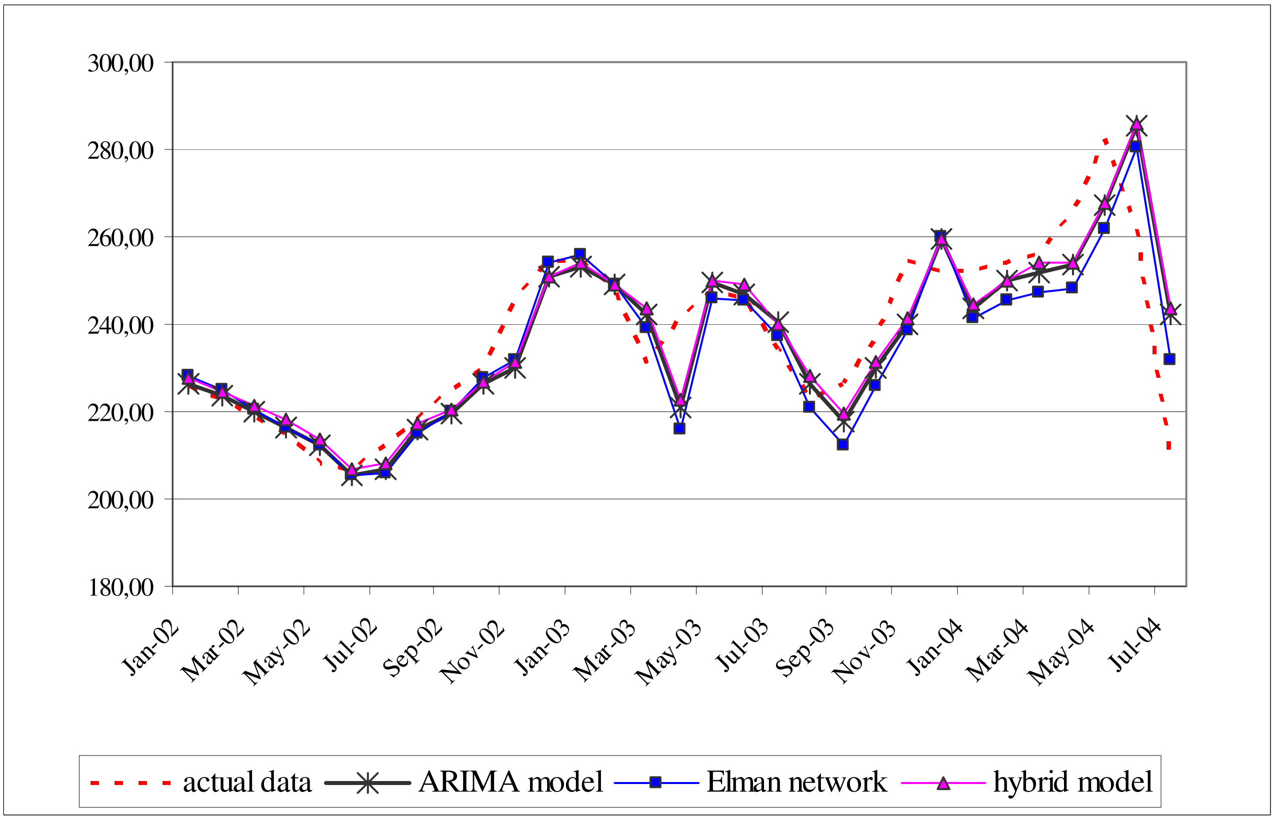

Fig. A 2: One-step-ahead forecasts of canola prices (Germany, January 2002 to July 2004)