Neurocomputing 275 (2018) 2831–2844

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China

Tao Xiong[a], Chongguang Li[a], Yukun Bao[b,*]

[a] College of Economics and Management, Huazhong Agricultural University, 430070 Wuhan, PR China
[b] School of Management, Huazhong University of Science and Technology, 430074 Wuhan, PR China

ARTICLE INFO

Article history:
Received 26 February 2016
Revised 24 November 2017
Accepted 26 November 2017
Available online 2 December 2017

Communicated by K Chan

Keywords:
Vegetable price forecasting
Seasonal forecasting
Seasonal adjustment
Seasonal-trend decomposition procedures based on loess (STL)
Extreme learning machines (ELM)

ABSTRACT

In view of the importance of seasonal forecasting of agricultural commodity price, particularly vegetable prices, and the limited research attention paid to it previously, this study proposes a novel hybrid method combining seasonal-trend decomposition procedures based on loess (STL) and extreme learning machines (ELMs) for short-, medium-, and long-term forecasting of seasonal vegetable prices. In the formulation of the proposed method (termed STL-ELM), the original vegetable price series are first decomposed into seasonal, trend, and remainder components. Then, the ELM is used to forecast the trend and remainder components independently, while the seasonal-naïve method is used to forecast seasonal components with a 12-month cycle. Finally, the prediction results of the three components are summed to produce an ensemble prediction of vegetable prices. In addition, an iterated strategy is used to implement multi-step-ahead forecasting. In terms of two accuracy measures and the Diebold-Mariano test, the experimental results show that the proposed method is the best-performing method relative to the competitors listed in this study, indicating that the proposed STL-ELM model is a promising method for vegetable price forecasting with high seasonality.

1. Introduction

According to *China Agricultural Yearbook 2011*,[1] the vegetable yield in China increased by 232% from 204 million tons in 1991 to 679 million tons in 2011, accounting for 52% of the total vegetable yield in the world. Both the yield and production value of the vegetables were greater than those of grain in 2011, indicating that vegetables have become one of the most important agricultural commodities in the daily life of Chinese people since 2011.

Vegetable prices have been exceptionally volatile over the past several years, reaching a low of 1.37 RMB/kg in July 2003 and a high of 11.73 RMB/kg in January 2014 (kidney bean is presented as an example here), as shown in Fig. 1. This has increased the risk faced by practitioners of agriculture when making decisions related to production, marketing, inventory, and resource allocation. In addition, the highly perishable nature of vegetables further intensifies this risk. Therefore, an understanding of price behavior is a critical element in making decisions under uncertain conditions that significantly affect the returns of the agricultural market participants. In this sense, most of the decisions are made at the beginning of the planting season, but not all of them are reversible. Therefore, price forecasting is a crucial step, and research on vegetable price forecasting is of great significance [1].

A significant characteristic of vegetable price series is the seasonality (as shown in Fig. 1), which is the biggest obstacle for obtaining accurate forecasts of vegetable prices. Given the complexity of the price series, many models have been specified for capturing the behavior of vegetable prices, but researchers have not reached a consensus on the best model for vegetable prices. The literature reviews on price forecasting in the agricultural commodity market, particularly vegetable markets, are detailed in Section 2.1.

An extensive investigation of the literature reveals that considerable efforts have been devoted to improving or developing techniques and tools to analyze and forecast time series with high seasonality. The seasonal auto-regressive integrated moving average (SARIMA) model [2] is the most popular and widely used technique for forecasting seasonal time series [3,4]. However, the forecasting performance may be very poor if the SARIMA built using linear assumptions for time series with high nonlinearity is used. Recently, the area of artificial neural networks (ANNs) has witnessed substantial improvements in terms of seasonal time series forecasting. Notable works on ANN in seasonal forecasting include those of Zhang and Qi [5] and Zhang and Kline [6]. More recently, support vector regression (SVR) has been extensively applied as

* Corresponding author.
E-mail addresses: yukunbao@hust.edu.cn, y.bao@ieee.org (Y. Bao).
[1] http://lib.cnki.net/cyfd/J149-N2012070010.html.

https://doi.org/10.1016/j.neucom.2017.11.053
0925-2312/

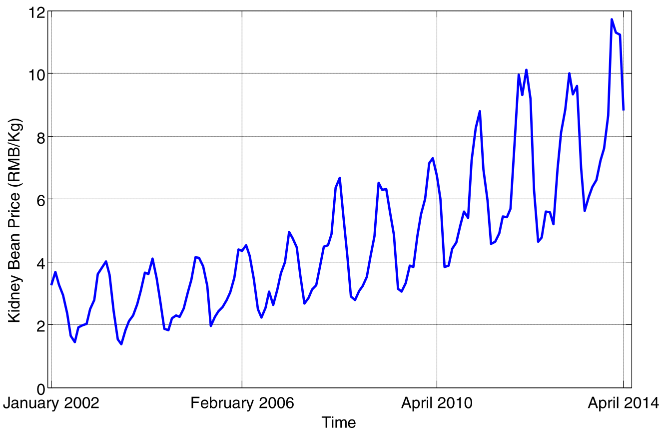


Fig. 1. Kidney bean price series.

a powerful alternative to traditional ANN for seasonal time series forecasting across a wide range of disciplines [7,8]. Although good performances using SVR for seasonal time series forecasting were reported in many areas, similar to the traditional ANN model, SVR also suffers from expensive computational costs and sensitivity to parameter selection.

Recently, another type of ANN, namely extreme learning machine (ELM), was proposed by Huang et al. [9]. ELM is an improved single-layer feed-forward neural network (SLFN) that uses a more efficient learning algorithm. The most attractive advantage of ELM is its extremely fast learning speed, and the generalization of ELM is superior to that of traditional ANN and comparable to that of SVR [9,10]. In addition, only the output weights of ELM need to be tuned relative to traditional ANN because the input weights and biases are randomly determined in advance in ELM [9]. Due to its excellent performance, till now, ELM has attracted more attention from the academic community and has been successfully used in many different areas, such as energy markets [11–13], environmental engineering [10], business management [14,15], medical science [16], and single processing [17]. To the best of our knowledge, however, the application of ELM for seasonal time series forecasting has not been widely explored.

In this study, the ELM, in conjunction with a decomposition technique, is selected to handle forecasting problems of seasonal vegetable price series. It has been documented in the literature that seasonal adjustment of time series using certain decomposition techniques before further forecasting is an efficient practice for improving seasonal time series forecasting. A well-established decomposition technique, seasonal-trend decomposition based on loess smoothing (STL) proposed by Cleveland et al. [18], is used for decomposition of the seasonal vegetable price series into its constituent components in this study. The most important advantages of the STL, relative to other decomposition techniques, are its strong robustness to outliers in the examined time series and ease of use [19]. In addition, the seasonal component obtained using STL always exhibits a rigorous cycle, which facilitates forecasts by means of the seasonal-naïve method. In this study, for example, the seasonal component of the vegetable price exhibits a standard 12-month cycle. Thus, the seasonal-naïve method is the optimal model for forecasting the seasonal component.

Motivated by the points mentioned above, a hybrid method combing STL and ELM, termed STL-ELM, is developed for the seasonal forecasting of vegetable price series in this study. In the proposed method, STL is first used to decompose the examined vegetable price series into seasonal, trend, and remainder components. Second, ELM is applied to forecast the trend and remainder components, while the seasonal-naïve method is used to forecast the seasonal component with a 12-month cycle. Finally, the prediction results of three components are combined by simple summing to generate an aggregated output, which can be seen as the final prediction for the original vegetable price series. The monthly prices of five vegetables, i.e., cabbage, pepper, cucumber, green bean, and tomato, from the Chinese agriculture market are chosen as experimental datasets. Apart from the individual ELM, for comparison purposes, four well-established methods used in seasonal time series forecasting, i.e., SARIMA [3], time delay neural network (TDNN) [20], SVR [7], and SARIMA-Kalman filter [21], are selected as benchmarks. The quantitative assessments are established on the basis of short-, medium-, and long-term forecasting performance in terms of two accuracy measures and the Diebold-Mariano test. In terms of the implementation of medium-term (three-step-head) and long-term (six-step-head) forecasting, the extensively used iterated strategy [22] is used in this study.

This paper is organized as follows. In Section 2, a literature review and the contributions of this paper are introduced. Section 3 presents a brief description of STL, ELM, and the iterated strategy, followed by a formulation of the proposed STL-ELM model in detail. The research design of the data description, measurement criteria, and benchmark prediction models is provided in Section 4. Section 5 illustrates and discusses the experimental results for five vegetable price series. Concluding remarks are finally made in Section 6.

2. Literature review and research contribution

2.1. Literature review

Although accurate forecasts of vegetable prices can provide critical and useful information for agricultural market participants, it is easy to see that studies on vegetable price forecasting are limited in number. Dieng [23] investigated the performance of three traditional parametric models for forecasting tomato, potato, and onion prices and further made recommendations to potential users in Senegal. To analyze the seasonal variations of tomato prices, Adanacioglu and Yercan [24] developed a SARIMA model to forecast the monthly tomato price at a wholesale level in Turkey. Martin-Rodriguez and Caceres-Hernandez [1] proposed a restricted evolving spline model to capture the dynamic process of change in the seasonal pattern of tomato prices and further forecast them. Taking advantage of computational intelligence, Nasira and Hemageetha [25] developed back-propagation neural network (BPNN)-based prediction models to forecast tomato prices. Amiri et al. [26] examined and compared the ability of geostatistical models, inverse distance weighting, adaptive neuro-fuzzy inference systems, and the winter method for seasonal forecasting of potato and onion prices in Iran. Melin et al. [27] developed several ANN-based prediction models for simulating and predicting the dynamic behavior of tomato prices in the U.S. Hemageetha and Nasira [28] developed a prediction model based on radial basis function neural networks to predict tomato prices.

In addition, some analyses and discussions on the existing literature on price forecasting related to agricultural commodity spot markets will help provide some useful suggestions and implications for seasonal forecasting of vegetable prices. Colino et al. [29] examined and compared the predictive performance of outlook hog price forecasts with several alternative time series methods. Davenport and Funk [30] proposed a characteristic-based clustering method to facilitate forecasts of cross-country grain price data and identify dissimilarities in price behavior across multiple markets in Kenya. Three ANN-based models and an integrated prediction model were developed by Luo et al. [31] to forecast the price of Lentinus edodes in Beijing. Li et al. [32] developed a short-term prediction model based on chaotic ANN for forecasting the weekly retail prices of eggs in China. After testifying that the seasonality present in cereal prices is deterministic, Jumah and Kunst [33] applied traditional in-sample fit and moving-window techniques to model the seasonal pattern and forecast the cereal prices. Jha and Sinha [20] examined and compared the price forecasting capabilities of ARIMA and time-delay neural network (TDNN) in forecasting prices of soybean, groundnut, rapeseed, and mustard in India. The results obtained in [20] show that the TDNN in general provides better forecasts as compared to ARIMA for nonlinear patterns. As such, TDNN is adopted as the benchmark in this study. To capture the mean dependence, volatility, and, more importantly, high frequency of zero changes in German milk-based commodities prices, Kömm and Küsters [34] proposed a Markov switching mixture model for forecasting milk-based commodity prices in Germany. Onour and Sergi [35] developed two competing models, i.e., the thin-tailed normal distribution and the fat-tailed student *t*-distribution models, to model and forecast the volatility of prices of wheat, rice, sugar, beef, coffee, and groundnut. Ribeiro and Oliveira [36] proposed a hybrid model integrating the ANN and Kalman filter techniques for forecasting the prices of sugar. Ramirez and Fadiga [37] examined the ability of asymmetric-error generalized autoregressive conditional heteroskedasticity (GARCH) models to forecast soybean, sorghum, and wheat prices in the U.S. Shih et al. [38] proposed a weighted case-based reasoning approach to construct a price prediction model for broiler prices. Ticlavilca [39] proposed a multi-step-ahead prediction model based on multivariate relevance vector machines to forecast the prices of cattle, hog, and corn in the U.S. By extracting two factors from a panel of commodity prices themselves using principal components analysis, West and Wong [40] constructed a static factor model to model and forecast the monthly prices of corn, cotton, and soybean. Zhang et al. [41] developed an aquatic product price forecasting support system. Zou et al. [42] compared the predictive ability of ARIMA, neural networks, and a hybrid model for forecasting wheat prices in China.

A summary of the models, forecasted agricultural commodity type, prediction horizon, seasonality, and time scale of these studies is provided in Table 1.

2.2. Research contribution

As seen in Table 1, in summary, our contributions can be outlined as follows.

(1) Because only seven out of 23 studies are associated with vegetable prices, the current study adds to a fairly limited body of research in this area. (2) Only seven studies paid attention to the seasonal pattern in agricultural commodity prices, but none of them investigated the ability of decomposition techniques to address seasonal variation. To this end, a well-established STL is applied to handle seasonal patterns of vegetable prices and further facilitate forecasts in this study. (3) The neural network is the most commonly used computational intelligence tool in these studies. However, neural networks suffer from expensive computational costs and sensitivity to parameter selection. Thus, a novel modeling technique, ELM, is used in the current study. A novel hybrid method combing STL and ELM is proposed for the seasonal forecasting of vegetable prices. (4) Of the 23 studies conducted, most focused on one-step-ahead forecasting, and only one study conducted multi-step-ahead forecasting for vegetable prices. Thus, the present study provides the first experimental evidence within the literature to examine the capability of computational intelligence methods and traditional statistical methods for short-, medium-, and long-term forecasting for vegetable prices. (5) In terms of the few studies focused on the Chinese market (only 4 studies focus on Chinese market), this study provides empirical evidence on agricultural commodity price forecasting in the Chinese market.

3. Methodologies

This section introduces the general process of the proposed STL-ELM method. First, the STL and ELM are briefly presented. Then, the proposed STL-ELM method is formulated, and the corresponding steps are described in detail.

3.1. Seasonal-trend decomposition procedure based on loess (STL)

The STL method, originally proposed by Cleveland et al. [18], is a filtering procedure for decomposing a time series into additive variation components. Significantly different from traditional seasonal decomposition techniques, such as X-12-ARIMA and the ratio-to-moving-average method, STL provides more robust results when dealing with the outliers in the examined time series.

Given a vegetable price series X_t, STL decomposes X_t into the three additive components of seasonal S_t, trend T_t, and remainder R_t components: $X_t = S_t + T_t + R_t$. STL is an iterative method consisting of two recursive procedures, inner and outer loops. Each of the passes through the inner loop consists of a seasonal smoothing that updates the seasonal component, followed by a trend smoothing that updates the trend component once. Upon the completion of the inner loop, robustness weights are calculated in the outer loop, which are then applied to decrease the influence of the outliers on updating seasonal and trend components in the following inner loop. Specifically, the inner loop is composed of six steps as follows:

Step 1: Detrending. At iteration $k+1$ of the inner loop, the original series X_t is detrended with the estimated trend component $T_t^{(k)}$ obtained at the kth pass: $X_t^{detrend} = X_t - T_t^{(k)}$.
Step 2: Seasonal smoothing. The sub-cycle series $X_t^{detrend}$ is smoothed by a Loess smoother to obtain a preliminary seasonal component $\tilde{S}_t^{(k+1)}$.
Step 3: Low-pass filtering of the smoothed seasonality. $\tilde{S}_t^{(k+1)}$ obtained in *Step 2* is processed using a low-pass filter, followed by a Loess smoother, to identify any remaining trend $\tilde{T}_t^{(k+1)}$.
Step 4: Detrending of the smoothed seasonality. The additive seasonal component $S_t^{(k+1)}$ is computed as the difference between the low-pass values and the preliminary seasonal component: $S_t^{(k+1)} = \tilde{S}_t^{(k+1)} - \tilde{T}_t^{(k+1)}$.
Step 5: Deseasonalizing. The original series X_t is reduced by the seasonal component $S_t^{(k+1)}$ to obtain a seasonally adjusted series $X_t^{deseason} = X_t - S_t^{(k+1)}$.
Step 6: Trend smoothing. The seasonally adjusted series $X_t^{deseason}$ obtained in *Step 5* is smoothed by a Loess smoother to obtain the trend component $T_t^{(k+1)}$:

In the outer loop, the trend and seasonal components obtained in the inner loop are used to compute the remaining component $R_t^{(k+1)} = X_t - S_t^{(k+1)} - T_t^{(k+1)}$. Any large values in $R_t$are treated as outliers and a weight is computed. In the next iteration of inner loop, the weight is employed to down-weight the effect of outliers identified in the previous iteration of the outer loop.

As such, the original series X_t is decomposed into the three additive components of seasonal S_t, trend T_t, and remainder R_t components using the STL technique.

3.2. Extreme learning machine (ELM)

The extreme learning machines (ELM), originally proposed by Huang et al. [9], is a novel learning algorithm for a single hidden-layer feed-forward neural network (SLFN). Significantly different

Table 1
Studies on price forecasting relating to agricultural commodity spot market.

Author	Models	Type	Prediction horizon (Prediction strategy)	Whether deal with seasonal	Time scale
Dieng (2008)	Naïve, exponential smoothing, and ARIMA	Tomato, potato, and onion	1	No	Monthly
Adanacioglu and Yercan (2012)	SARIMA	Tomato	4 (Iterative Strategy)	SARIMA	Monthly
Martin-Rodriguez (2013)	Restricted evolving spline model	Tomato	1	Spline functions	Daily
Nasira and Hemageetha (2012)	BP neural network	Tomato	1	No	Weekly
Amiri et al. (2011)	Geostatistical models, neuro fuzzy approach, and Winter method	Potato and onion	1	No	Quarterly
Melin et al. (2007)	Modular neural networks	Tomato	1	No	Monthly
Hemageetha and Nasira (2013)	BP neural network, RBF neural network	Tomato	1	No	Weekly
Colino et al. (2011)	VARs, Bayesian VARs	Hog	1, 2, and 3 (-)	Seasonal dummy variables	Quarterly
Davenport and Funk (2015)	Characteristic based clustering-based prediction model	Grain	1, 2, and 3 (Iterated strategy)	SARIMA	Monthly
Luo et al. (2011)	BP neural network, RBF neural network	Lentinus edodes	1	No	Monthly
Li et al. (2013)	Chaotic neural network	Egg	1,2,3,4, and 5 (-)	No	Weekly
Jumah and Kunst (2008)	Restricted rank model and unrestricted rank model	Barley and wheat.	1,2,..., 19, and10 (-)	Seasonal dummy variables	Quarterly
Jha and Sinha (2014)	Time-delay neural networks (TDNN)	Oilseeds	1,3,6, and 12 (Iterative Strategy)	Seasonality-adjusted method	Monthly
Kömm and Küsters (2015)	Zero-inflated models, ARIMA(1,1,0)-GARCH(1,1) model	Milk-based Commodities	1,2,3,4, and 5 (-)	No	Weekly
Onour and Sergi (2011)	Thin tailed the normal distribution, fat-tailed Student t-distribution models	Wheat, rice, beef, groundnut, sugar, and coffee	1	No	Monthly
Ribeiro and Oliveira (2011)	Hybrid model based on neural networks and Kalman filter	Sugar.	1	No	Monthly
Ramirez and Fadiga (2003)	GARCH	Soybean, sorghum, and wheat	1	Centered multiplicative moving-average procedure	Quarterly
Shih et al. (2009)	Case-based reasoning approach	Broiler	1	No	Monthly
Ticlavilca et al. (2010)	Multivariate relevance vector machine	Cattle, hog, and corn	1, 2, and 3 (Multi-input Multi-output strategy)	No	Monthly
West and Wong (2014)	The factor model based on principal components extracted from commodity prices themselves	Corn, cotton, soybean meal, soybean oil, and wheat	1,3, and 12 (Direct strategy)	No	Monthly
Zhang et al. (2005)	Moving average, linear regression, and neural networks	Aquatic product	1	No	Annual, quarterly, monthly, weekly, and 3-daily
Zou et al. (2007)	Hybrid model based on ARIMA and neural networks	Wheat	1	No	Monthly

- not reported/unclear.

from the gradient-based learning algorithm used for a traditional ANN, the input weights and biases are randomly determined, and subsequently, the output weights are tuned using simple matrix computations in ELM, dramatically saving training time.

Given a set of N distinct samples $\{(\mathbf{x}_j, \mathbf{t}_j)\}_{j=1}^{N}$ with inputs $\mathbf{x}_j = [x_{j1}, x_{j2}, \dots, x_{jn}]^T \in \Re^n$ and outputs $\mathbf{t}_j = [t_{j1}, t_{j2}, \dots, t_{jm}]^T \in \Re^m$, the ELM with $\tilde{N}$ hidden neurons and activation function $g(\cdot)$ is mathematically modeled as

$$\sum_{i=1}^{\tilde{N}} \beta_i g\big(\mathbf{w}_i \cdot \mathbf{x}_j + b_i\big) = \mathbf{t}_j, \quad j = 1, 2, \dots, N \tag{1}$$

where $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ denotes the weight vector connecting the input neurons and the ith hidden neuron, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ denotes the weight vector of the connection from the ith hidden neuron to the output neuron, and b_i denotes the threshold of the ith hidden node.

Eq. (1) can be rewritten compactly as $\mathbf{H}\beta = \mathbf{T}$, where $\beta = [\beta_1, \dots, \beta_{\tilde{N}}]^T$, $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_N]^T$ and $\mathbf{H}$ is the hidden layer output matrix of ELM, expressed as

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \dots & g\big(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}\big) \\ \vdots & \dots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \dots & g\big(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}\big) \end{bmatrix} \tag{2}$$

In ELM, the input weights $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ and hidden biases b_i are randomly assigned and do not, in fact, require any tuning in the following training process. Accordingly, calculating the output weights $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is equivalent to simply finding the least square (LS) solution to the above linear system $\mathbf{H}\beta = \mathbf{T}$. The minimum norm LS solution of Eq. (1) is $\hat{\beta} = \mathbf{H}^{\dagger}\mathbf{T}$, where $\mathbf{H}\dagger$ is the Moore–Penrose (MP) generalized inverse of matrix $\mathbf{H}$.

3.3. Iterated strategy for multi-step-ahead forecasting

The prediction strategy is a fundamental issue that must be addressed when establishing models for multi-step-ahead forecasting. The most common strategy is what Chevillon [22] in a recent survey calls the iterated strategy, which has been widely used in the forecasting community [20,43–45].

An iterated strategy constructs a prediction model by minimizing the squares of the in-sample one-step-ahead residuals and then uses the predicted value as an input for the same model to forecast the subsequent point, continuing in this manner until reaching the horizon. The main advantage of the iterated strategy relative to other strategies is that only one model must be built, and thus, it has low computational cost. The formulation of the iterated strategy is briefly presented as follows.

The iterated prediction strategy learns the one-step-ahead prediction model:

$$x_{i+1} = f(x_i, \ldots, x_{i-d+1}) + \varepsilon \tag{3}$$

where d is the maximum embedding order and ε is the scalar zero-mean noise term.

After the learning process, the estimation of the next H values is returned using:

$$\hat{x}_{i+h} = \begin{cases} \hat{f}(x_i, x_{i-1}, \ldots, x_{i-d+1}) & if \quad h = 1 \\ \hat{f}(\hat{x}_{i+h-1}, \ldots, \hat{x}_{i+1}, x_i, \ldots, x_{i-d+h}) & if \quad h \in [2, \ldots, d] \\ \hat{f}(\hat{x}_{i+h-1}, \ldots, \hat{x}_{i+h-d}) & if \ h \in [d+1, \ldots, H] \end{cases} \tag{4}$$

In this study, once the original vegetable price is decomposed into three components, the ELM coupled with the iterated strategy is used to model and forecast the trend and remainder components, which are detailed in the following subsection.

3.4. The proposed hybrid STL-ELM method

In this subsection, the proposed STL-ELM method is formulated, and the corresponding steps are presented in detail.

Given a vegetable price series X_t for $t = 1, 2, \ldots, n$, a three-step modeling procedure for the proposed STL-ELM method can be formulated for vegetable price forecasting (Fig. 2). As shown in Fig. 2 (kidney bean is presented as an example), the proposed STL-ELM method is generally composed of the following three main steps:

Step 1: Decomposition. The original kidney bean price series X_t is decomposed into the seasonal S_t, trend T_t, and remainder R_t components using the STL technique as follows: $X_t = S_t + T_t + R_t$, $t = 1, 2, \ldots, n$. STL is implemented using the *stl*[2] function of the library *stats* in the R statistical software environment.

Step 2: Single forecast. ELM is used as a single forecasting technique to independently model and forecast the extracted trend component T_t, $t = 1, 2, \ldots, n$ and remainder component R_t, $t = 1, 2, \ldots, n$ in a multi-step-ahead fashion using the iterated strategy (prediction horizon $H = 1$, 3, and 6);this implementation of ELM and the concept of multi-step-ahead forecasting are detailed as follows.

To demonstrate the concept of multi-step-ahead forecasting of ELM used in this study, the remainder component R_t, $t = 1, 2, \ldots, n$ is presented as an example. Given the remainder component R_t, $t = 1, 2, \ldots, n$, as discussed in Section 3.3, an ELM model is first trained by minimizing the squares of the in-sample one-step-ahead residuals. Fig. 3 illustrates the concept of multi-step-ahead forecasting with the obtained ELM model. In Fig. 3(a), using the historical values of the remainder component $(R_{t-d+1}, \ldots, R_{t-1}, R_t)$, the subsequent (one-step-ahead) value $\hat{R}_{t+1}$ is predicted. In Fig. 3(b), based on the historical values of the remainder component $(R_{t-d+2}, \ldots, R_{t-1}, R_t)$ and the previously predicted value $\hat{R}_{t+1}$ as the predictor, the two-step-ahead predicted value $\hat{R}_{t+2}$ is obtained. In Fig. 3(c), using the historical values of the remainder component $(R_{t-d+3}, \ldots, R_{t-1}, R_t)$ and the previously predicted values $\hat{R}_{t+1}$ and $\hat{R}_{t+2}$ as the predictors, the three-step-ahead predicted value $\hat{R}_{t+3}$ is obtained. Similarly, the six-step-ahead predicted value $\hat{R}_{t+6}$ is obtained.

Regarding the seasonal component, it is clear that S_t, $t = 1, 2, \ldots, n$ exhibits a standard 12-month cycle, i.e., $S_{t+12} = S_t$. Accordingly, the seasonal-naïve method is the optimal model for forecasting the seasonal component. In the seasonal-naïve method, the future value at period $t + 12$ is equivalent to the historical value at period t. This model can be represented as follows:

$$\hat{S}_{t+12} = S_t,\ t = 1, 2, \ldots, n \tag{5}$$

where S_t is the true value at period t of the seasonal component and $\hat{S}_{t+12}$ is the predicted value at period $t + 12$of the seasonal component.

Step 3: Summation. The prediction results of the trend and remainder components generated by ELM and the seasonal component using the above seasonal-naïve method in *Step 2* are summed to produce an aggregated output, which is the final predicted price of the original kidney bean price.

4. Research design

In this section, the research design of the data description, performance measurement criteria, and benchmark prediction models are provided in detail. The next section reports additional experimental results and the discussion.

4.1. Data description

The monthly price series of five vegetables, i.e., cabbage, pepper, cucumber, green bean, and tomato, are used as experimental datasets in this study because these vegetables play an important role in the daily life of Chinese people. The datasets are collected from Chinese bazaars and are freely available from the *China Monthly Statistics* provided by the National Bureau of Statistics of the People's Republic of China.[3]

The objective of the current study is to propose an STL-ELM method for seasonal vegetable price forecasting with various prediction horizons ($H = 1$, 3, and 6). The monthly prices of five vegetable price series from January 2002 to April 2014 are used here, depending on the availability. For each vegetable, the price series are split into the estimation and holdout samples following the commonly used two-thirds norm. Accordingly, the first 98 observations from January 2002 to February 2010 are selected as the estimation sample, which is used to determine the unknown parameters of the pre-defined models. The last 50 observations from

[2] Source code is available at http://web.mit.edu/people/jhaas/MacData/afs/sipb/project/r-project/arch/sun4x_59/lib/R/library/stats/html/stl.html.

[3] The datasets are available at http://www.stats.gov.cn/.

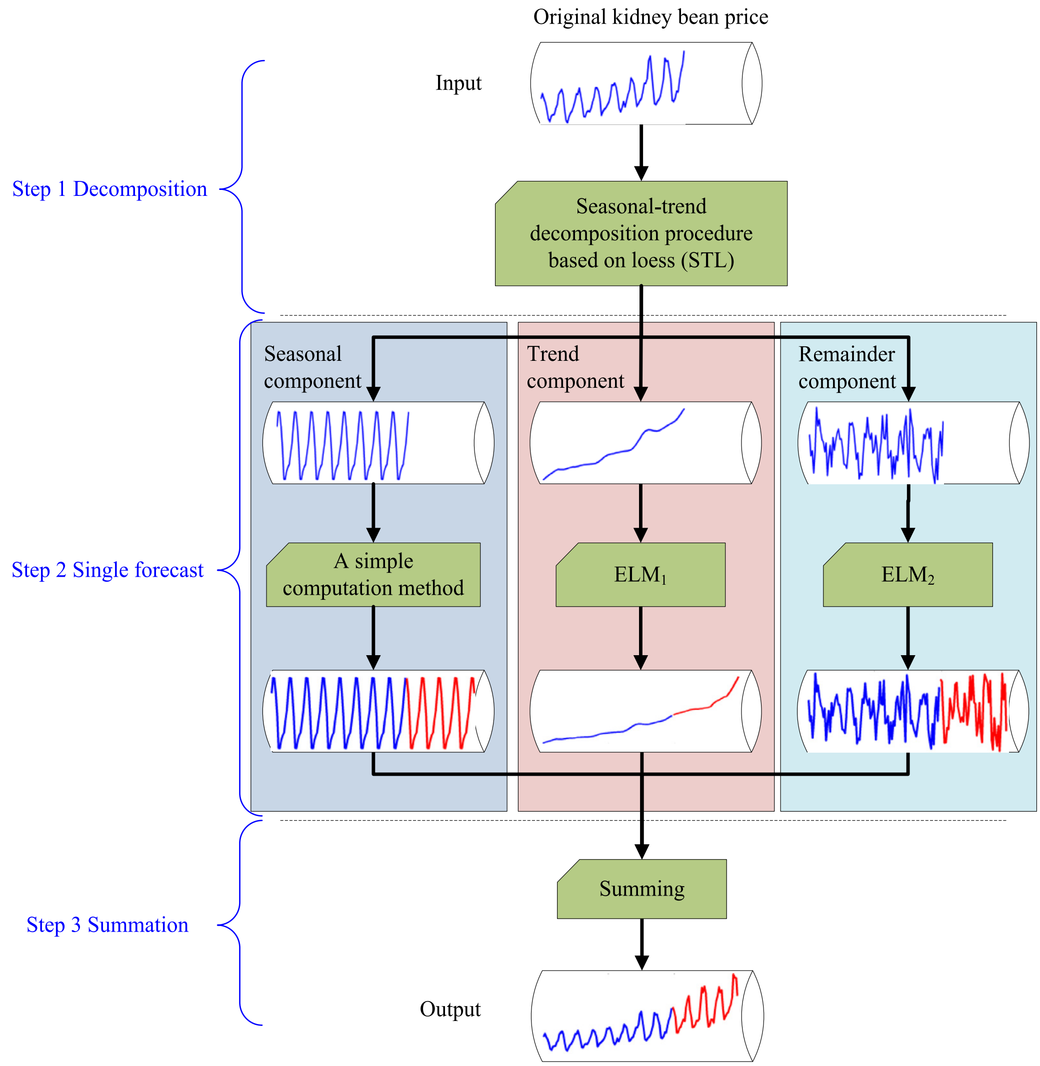


Fig. 2. The proposed STL-ELM model for vegetable price forecasting.

March 2010 to April 2014 are used as the holdout sample, which is applied to evaluate the forecasting performance of the obtained models. To evaluate the prediction ability of the proposed STL-ELM method, we carry out one-, three-, and six-step-ahead forecasting to examine the short-, medium-, and long-term forecasting ability, respectively. It should be noted that an iterated strategy [22,43,44] is used for implementing the medium- and long-term forecasting. In addition, in order to avoid inputs in greater numeric ranges from dominating those in smaller numeric ranges, a linear transformation is adopted to adjust the original vegetables prices series scaled into the range of [0, 1] in this study.

Fig. 4 depicts the monthly price series mentioned above. From Fig. 4. it can be seen that all five vegetable price series have a significant cyclical characteristic, which is mainly determined by the seasonality of vegetable production and consumption. In addition, a growth trend can be observed in these price series, which is mainly influenced by the inflation.

4.2. Measurement criteria

To compare the forecasting performance of the proposed STL-ELM method with that of four selected competitors, no single accuracy measure can capture the distributional features of the errors. Accordingly, two indices, the symmetric mean absolute percentage error (SMAPE) and mean absolute scaled error (MASE), are used as prediction accuracy measures. These indices are as follows:

$$\mathrm{SMAPE} = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{x_t - \hat{x}_t}{\left(|x_t| + |\hat{x}_t|\right)/2}\right| \tag{6}$$

$$\mathrm{MASE} = \frac{1}{T}\sum_{t=1}^{T}\left(\left|\frac{x_t - \hat{x}_t}{\frac{1}{N-1}\sum_{i=2}^{N}|x_i - x_{i-1}|}\right|\right) \tag{7}$$

where x_t is the observation at period t, $\hat{x}_t$ is the forecast of x_t, and T and N are the numbers of observations in the estimation sample and hold-out sample (in this case, $N = 98$ and $T = 50$), respectively.

Apart from the accuracy measures, the Diebold-Mariano (DM) statistic is used to test the statistical significance of two competing prediction models [46]. The DM test aims to test the null hypothesis of equality of the expected forecast accuracy against the alternative of different forecasting abilities across models. The loss function is set to the mean square prediction error (MSPE), and the null hypothesis is that the MSPE of the tested model *te* is not less than that of the reference model *re*. In particular, the DM statistic

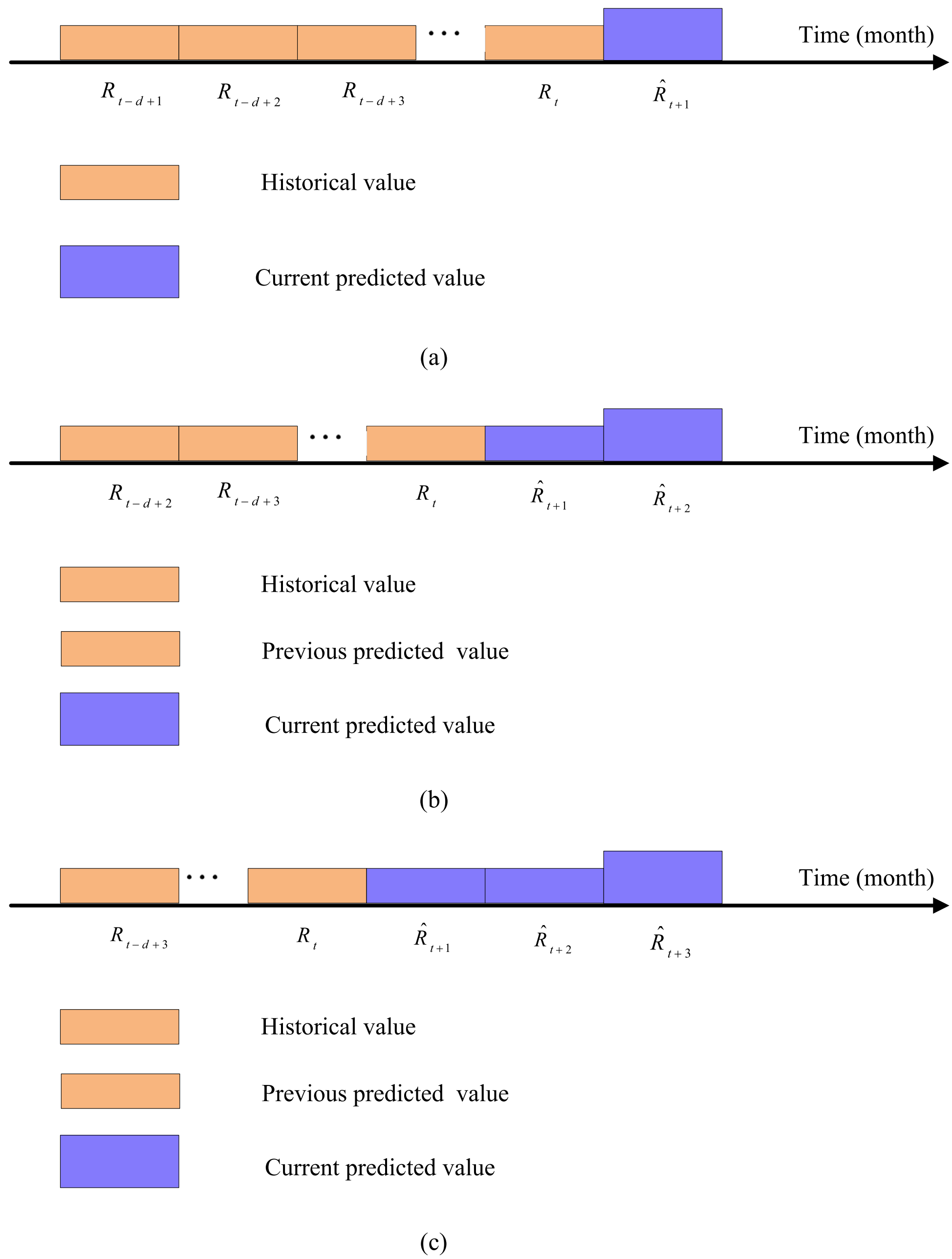


Fig. 3. Multi-step-ahead forecasting of ELM using iterated strategy. (a) one-step-ahead prediction. (b) two-step-ahead prediction. (c) three-step-ahead prediction.

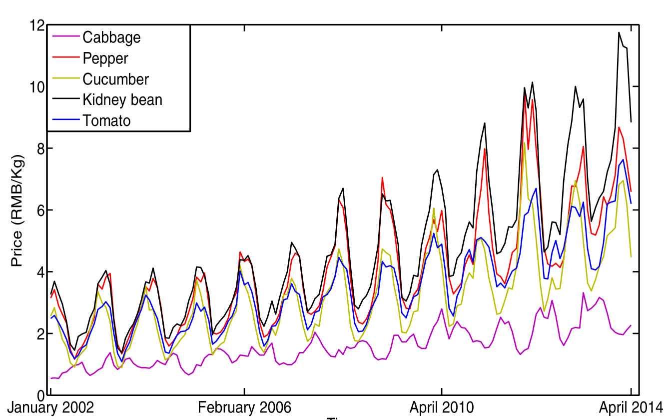


Fig. 4. Five vegetables price series.

can be defined as

$$S_{DM} = \frac{\bar{g}}{\sqrt{\left(\widehat{V}_g/N\right)}} \tag{8}$$

where $\bar{g} = (\sum_{t=1}^{N} g_t)/N$, $g_t = (x_t - \hat{x}_{te,t})^2 - (x_t - \hat{x}_{re,t})^2$, and $\widehat{V}_g = \gamma_0 + 2\sum_{t=1}^{\infty} \gamma_t$ $(\gamma_t = \mathrm{cov}(g_{t+1}, g_t))$. γ_0 is the variance of g_t, $\hat{x}_{te,t}$ and $\hat{x}_{re,t}$ represent the predicted values of x_t calculated using the tested method *te* and reference method *re*, respectively, in period *t*. *N* is the number of observations in hold-out sample.

4.3. Methodologies implementation

The purpose of this study is to propose a novel hybrid method combining STL and ELM for short-, medium-, and long-term

Table 2
The optimal number of hidden nodes in STL-ELM and pure ELM.

The number of hidden nodes	Cabbage	Pepper	Cucumber	Green bean	Tomato
The STL-ELM for trend component	8	6	7	6	8
The STL-ELM for reminder component	8	7	10	8	11
Pure ELM	7	9	7	10	7

forecasting of seasonal vegetable prices. The implementation of the proposed method is detailed as follows.

In this study, STL is implemented using the *stl* function of the library *stats* in the R statistical software environment. To execute the *stl* function, six parameters are to be determined in advance: (1) the number of observations n_p in each cycle, (2) the number of iterations of the inner loop n_i, (3) the number of iterations of the outer loop n_o, (4) the smoothing parameter n_s used for obtaining the seasonal component in *Step 2*, (5) the smoothing parameter n_l used for the low-pass filter in *Step 3*, and (6) the smoothing parameter n_t used for obtaining the trend component in *Step 6*. As such, six parameters in STL should be defined in advance. By doing so, the parameters in STL are determined according to the recommendations in [18,47] as follows: $n_p = 12$, $n_i = 1$, $n_o = 6$, $n_s = 12$, $n_l = 13$, and $n_t = 21$.

ELM is implemented using the *ELM package*[4] in the MATLAB computing environment. The number of input nodes is determined by the input selection using the filter method with partial mutual information. Specifically, the maximum embedding order d is set to 24 here. The number of output nodes is set as one, as the iterated strategy is used for implementing the multi-step-ahead forecasting. A logistic sigmoid function is selected as the activation function. As discussed in Section 3.2, ELM randomly determines the input weights and hidden biases, which do not require any tuning in the following training process. Thus, once these parameters are randomly generated in advance, the number of hidden nodes is determined in a trial-error fashion. We thereby construct twelve ELM models with various numbers of hidden nodes (varying from 4 to 15). Each ELM model is trained repeatedly 30 times on the estimation samples, and then, the average MSE of each ELM is calculated on the holdout sample. The ELM that yields the smallest average MSE value is selected as the best model. By doing so, the optimal number of hidden nodes in STL-ELM and pure ELM for five vegetable prices are shown in Table 2.

As discussed in Section 1.2, SARIMA, TDNN, SVR, ELM, and SARIMA-Kalman filter are selected as the benchmarks for the purpose of comparison. This subsection presents the implementation of these benchmarks in detail following two aspects, the software or program packages involved and parameter selection. The implementation of ELM, which is already presented above, is not repeated here.

(1) SARIMA: A vegetable price series X_t is SARIMA $(p, d, q) \times (P, D, Q)_S$ process with seasonal period length S if d and D are nonnegative integers and if the differenced series $W_t(1-B)^d(1-B^S)^D X_t$ is an ARIMA process, where B is the backshift operator defined by $B^a W_t = W_{t-a}$. In view of the strong seasonal pattern observed in five vegetable prices series, SARIMA is used as a benchmark here. The SARIMA is estimated using an automatic model selection algorithm implemented using the program package "forecast"[5] in the R statistical software environment. The optimal SARIMA models parameters for five vegetables prices are shown in Table 3.

(2) TDNN: To perform time series forecasting, the short-term memory existed in time series itself is required to make neural network dynamic. One simple way of building short-memory into the neural network is through the use of time delay. As such, the time delay neural network (TDNN) with one hidden layer used in this study is the commonly-used feed-forward neural network where lagged observations of the time series are used as the inputs in the input layer. Specifically, the number of input nodes which are lagged observations is determined using filter method with partial mutual information (maximum embedding order $d = 24$). The number of output nodes is set to one, and the number of hidden nodes is determined in a trial-error fashion. Similarly to optimization procedure for the number of hidden nodes in ELM presented above, we varied the number of hidden nodes from 4 to 15 with fivefold cross-validation method. The optimal number of input nodes, hidden nodes, and output nodes in TDNN for five vegetable prices are shown in Table 4. The logistic sigmoid function is also selected as the activation function. Different from the ELM model, the Levenberg and Marquardt Algorithm is used for training the TDNN. The TDNN is implemented using the *neural network toolbox* in the MATLAB computing environment.

Table 3
SARIMA model parameters for five vegetables prices.

SARIMA	Cabbage	Pepper	Cucumber	Green bean	Tomato
p	0	2	2	1	2
d	1	0	0	1	0
q	3	4	5	3	4
P	0	0	0	0	0
D	0	1	1	1	1
Q	3	2	2	3	2
S	12	12	12	12	12

(3) SVR: The SVR is implemented using *LibSVM* (version 2.86)[6] in the MATLAB computing environment. Analogously, the input variables of SVR are also determined by input selection using the filter method with partial mutual information (maximum embedding order $d = 24$). The widely used radial basis function (RBF) is employed as the kernel function for SVR modeling. The hyper-parameters (i.e., C, ε, and γ) are determined by the well-established and straightforward grid search method in a trial-error fashion.

(4) SARIMA-KF (Kalman filter): Although the SARIMA process is well-suited to capture seasonal pattern existed in vegetable prices, another limitation of the SARIMA is the difficulty of adjusting the model's parameters when new information arrives. By addressing this issue, Kalman filter (KF), proposed by Kalman [48], has been used successfully by several researchers [21,49] to tune the parameters of SARIMA or ARIMA model. Therefore, the SARIMA-KF method is also adopted as the benchmark here. For the detail formulation of pure KF and the hybrid SARIMA-KF method, please refer to [21,48,49]. In this study, after initializing the state equation and measurement equation for each vegetable price series based on the obtained SARIMA $(p, d, q) \times (P, D, Q)_S$ model presented in Table 3, the KF is used for vegetable price forecasting using a KF recursive step, such the one found in [21]. The KF is

[4] Source code is available at http://www.ntu.edu.sg/home/egbhuang/elm_codes.html.

[5] Source code is available at http://ftp.ctex.org/mirrors/CRAN/web/packages/forecast/index.html.

[6] Source code are available at http://www.csie.ntu.edu.tw/~cjlin/libsvm/.

Table 4
The final number of input, hidden, and output nodes in TDNN.

	Cabbage	Pepper	Cucumber	Green bean	Tomato
The number of input nodes	12	10	9	10	11
The number of hidden nodes	8	10	8	6	8
The number of output nodes	1	1	1	1	1

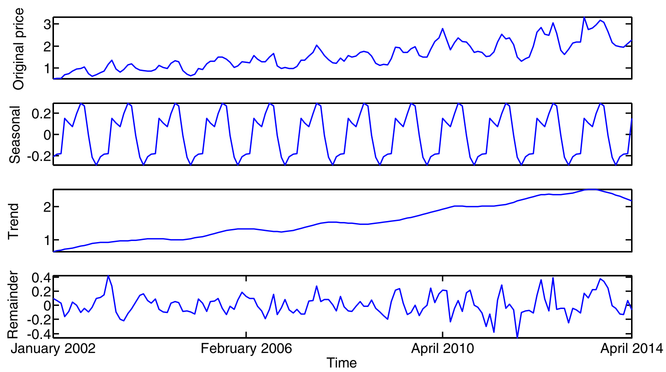


Fig. 5. The decomposition results of cabbage price series via STL.

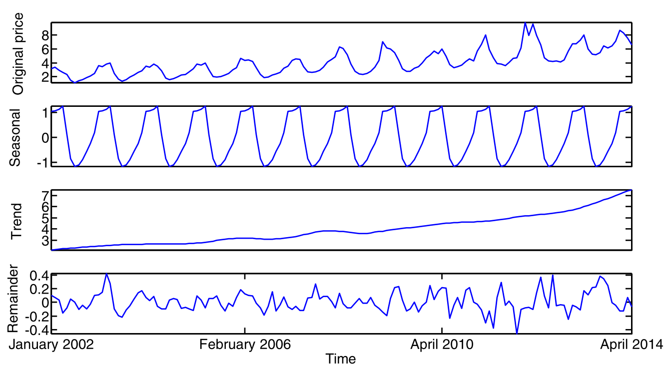


Fig. 6. The decomposition results of pepper price series via STL.

implemented using *toolbox*[7] in the MATLAB computing environment in this study.

5. Results and discussions

Using the research design mentioned above, the forecasting experiments for the vegetable price series are conducted. Accordingly, the forecasting performances of all of the examined models are evaluated using the two accuracy measures and the Diebold-Mariano test.

The decomposition results of the five vegetable price series using STL are shown in Figs. 5–9. It can be seen that each original vegetable price series is decomposed into seasonal, trend, and remainder components using the STL technique. All of the seasonal components of these vegetable price series show a 12-month cycle. In addition, a constant growth trend can be seen in all of the trend components, except for cabbage, where a downward trend can be observed after 2014. After the decomposition, as discussed in Section 3.4, the ELM coupled with the iterated strategy is used to forecast the extracted trend and remainder components, while a seasonal-naïve method is used to forecast the extracted seasonal component, as shown in Eq. (5). Finally, the prediction results of

[7] Source code are available at http://www.cs.ubc.ca/~murphyk/Software/Kalman/kalman.html#other.

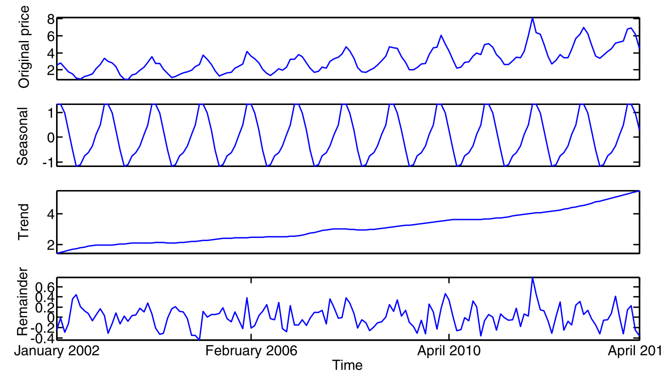


Fig. 7. The decomposition results of cucumber price series via STL.

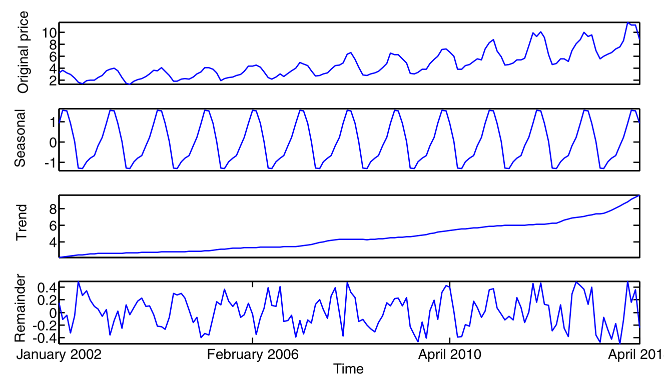


Fig. 8. The decomposition results of kidney bean price series via STL.

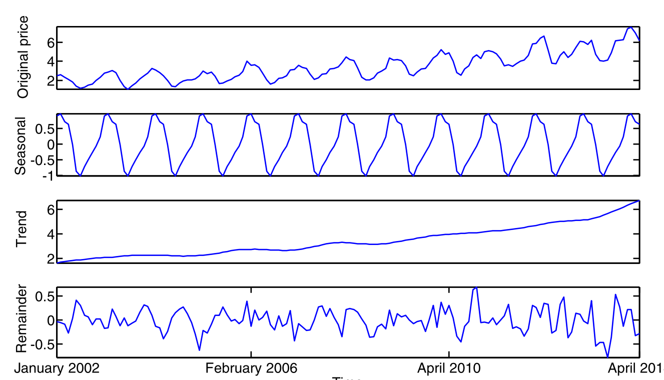


Fig. 9. The decomposition results of tomato price series via STL.

the seasonal, trend, and remainder components are summed to generate an aggregated output.

The forecasting performances of the six examined models (i.e., STL-ELM, ELM, SARIMA-KF, SVR, TDNN, and SARIMA) across the three prediction horizons (i.e., $H = 1$, 3, and 6) for SMAPE and MASE are shown in Table 5. In each row of Table 5 below, the entry with the smallest value is set in boldface.

From Table 5, it is clear that the proposed STL-ELM method is the best one for seasonal vegetable price forecasting across all pre-

Table 5
Prediction accuracy measures of different models on hold-out sample.

Accuracy measure	Prediction horizon (H)	Models					
		STL-ELM	ELM	SARIMA-KF	TDNN	SVR	SARIMA
Panel A: Cabbage price							
SMAPE	1	**3.217**	4.013	4.127	4.081	4.185	5.421
	3	**3.109**	4.118	4.002	4.140	4.106	5.854
	6	**3.584**	3.724	3.954	4.217	4.659	5.795
MASE	1	**0.847**	0.896	0.871	0.893	0.901	1.185
	3	**0.850**	0.903	0.896	0.921	0.914	1.312
	6	0.948	**0.922**	0.974	0.992	1.085	1.385
Panel B: Pepper price							
SMAPE	1	**3.231**	4.157	3.895	4.085	4.105	5.487
	3	**3.214**	4.214	3.931	4.384	4.245	5.285
	6	**3.715**	4.951	4.495	4.795	4.864	5.712
MASE	1	0.829	0.879	**0.814**	0.846	0.881	1.076
	3	**0.837**	0.918	0.912	0.934	0.907	1.105
	6	**0.886**	0.961	0.951	0.962	0.955	1.157
Panel C: Cucumber							
SMAPE	1	**3.197**	3.952	3.817	3.924	3.894	5.374
	3	**3.214**	4.128	3.954	4.172	4.084	5.814
	6	**3.469**	4.765	4.120	3.827	3.774	5.915
MASE	1	**0.840**	0.894	0.881	0.884	0.885	1.157
	3	**0.847**	0.894	0.891	0.901	0.886	1.285
	6	0.984	1.185	1.084	1.017	**0.978**	1.402
Panel D: Green bean price							
SMAPE	1	**3.185**	3.986	3.824	4.141	4.054	5.385
	3	**3.118**	3.895	3.924	3.876	3.968	5.796
	6	3.695	**3.459**	4.015	4.217	4.596	5.693
MASE	1	**0.839**	0.887	0.864	0.869	0.895	1.053
	3	**0.859**	0.915	0.894	0.923	0.908	1.284
	6	**0.918**	0.934	0.938	1.093	1.108	1.345
Panel E: Tomato price							
SMAPE	1	**3.205**	3.968	3.758	3.934	4.085	5.328
	3	**3.354**	4.168	3.958	4.281	4.129	5.285
	6	**3.585**	4.851	4.594	4.829	4.749	5.914
MASE	1	0.851	0.879	**0.842**	0.902	0.886	1.204
	3	**0.865**	0.914	0.896	1.054	0.909	1.285
	6	**0.924**	1.185	1.018	1.106	1.058	1.685

Note: For each row of the table, the entry with the smallest value is set in boldface.

diction horizons (i.e., one-, three-, and six-step-ahead) for the five vegetable price series, relative to the other five competitors listed in this study, though there are a few exceptions. It is conceivable that the reason behind the inferiority of the ELM, TDNN, and SVR relative to STL-ELM is that these three pure techniques cannot model seasonality directly, keeping in line with the works of Zhang and Qi [5] and Hong [7]. Therefore, the prior data processing, such as decomposition and seasonal adjustment, is necessary and critical to build a better forecaster [5], which is implemented as the proposed STL-ELM method in this study.

In addition, from all of the models examined in this study, the SARIMA is consistently the worst forecasting model for all of the vegetable price series, regardless of the prediction horizons and accuracy measures considered. It is conceivable that the reason behind the inferiority of the SARIMA is that it is a class of typical linear model and thus cannot capture nonlinear patterns hidden in the vegetable prices.

Apart from the proposed STL-ELM and SARIMA models, which perform the best and worst, respectively, among all of the examined models, other listed models produce some interesting mix results, which are analyzed in terms of short-term ($H=1$), medium-term ($H=3$), and long-term ($H=6$) forecasting.

First, regarding short-term ($H=1$) forecasting, the results of the SMAPE and MASE criteria show that STL-ELM performs the best, followed by SARIMA-KF, ELM, TDNN, and SVR. The worst model is the SARIMA model. The SMAPE values for STL-ELM are 3.217, 3.231, 3.197, 3.185, and 3.205 for cabbage, pepper, cucumber, green bean, and tomato price forecasting, respectively, which is clearly less than those of the other listed models. In the pure artificial intelligence (AI) models (i.e., ELM, TDNN, and SVR), it can be seen that ELM performs slightly better than TDNN. One exception occurs in the case of the cabbage price with SMAPE where the TDNN slightly outperforms the ELM. In addition, ELM and TDNN perform better than SVR, though there are a few exceptions.

Second, with respect to medium-term ($H=3$) forecasting, SARIMA-KF also ranks second after STL-ELM, followed by the SVR, TDNN, ELM, and SARIMA models in terms of SMAPE and MASE. The reason behind the superiority of the SARIMA-KF relative to the four pure models is due to the use of a Kalman filter to handle stochastic uncertainty of the vegetable price [21]. The MASE values of STL-ELM are 0.850, 0.837, 0.847, 0.859, and 0.865 for cabbage, pepper, cucumber, green bean, and tomato price forecasting, respectively, which are the minimum values. In addition, all of the MASE values of SARIMA are close to 1.350, which is the maximum value. Different from the ranking among ELM, SVR, and TDNN for short-term forecasting, SVR is superior to ELM and TDNN in most cases, regardless of the accuracy measures and vegetable price series considered, though there are a few exceptions.

Third, in the case of long-term ($H=6$) forecasting, some similar observations can be drawn. The proposed STL-ELM method performs better than all of the other competitors for all of the cases, though there are a few exceptions, followed by SARIMA-KF, ELM, SVR and TDNN, almost with a tie, and SARIMA. Three exceptions are observed when considering the cabbage price with MASE and the green bean price with SMAPE, where the ELM slightly outperforms the STL-ELM, and when considering the cucumber price with MASE, where SVR slightly outperforms the STL-ELM. In all cases, the SARIMA invariably delivers the worst performance.

Table 6
DM test results for different models for one-step-ahead forecasting.

Vegetable type	Tested model	Reference model				
		ELM	SARIMA-KF	TDNN	SVR	SARIMA
Cabbage	STL-ELM	−1.46 *	−1.50 *	−1.48*	−1.42 *	−4.78***
	ELM		−0.28	−0.33	−0.31	−3.10***
	SARIMA-KF			0.13	−0.49	−3.12 ***
	TDNN				−0.43	−3.25***
	SVR					−3.08***
Pepper	STL-ELM	−1.68**	−1.52*	−1.88**	−1.65**	−4.21***
	ELM		0.34	0.09	−0.08	−3.79***
	SARIMA-KF			−0.37	−0.30	−4.19***
	TDNN				0.24	−3.24***
	SVR					−3.32***
Cucumber	STL-ELM	−2.14 **	−2.02**	−2.24**	−2.16**	−3.57 ***
	ELM		−0.35	0.29	0.12	−3.28 ***
	SARIMA-KF			−0.75	−0.92	−3.35***
	TDNN				−0.69	−3.52 ***
	SVR					−3.64***
Green bean	STL-ELM	−1.65 **	−1.53*	−1.75**	−1.69 **	−4.81***
	ELM		0.51	−0.21	−0.39	−3.67 ***
	SARIMA-KF			0.48	0.56	−3.74***
	TDNN				0.39	−3.85 ***
	SVR					−3.72 ***
Tomato	STL-ELM	−1.97 **	−1.78**	−1.65**	−1.42 *	−4.28***
	ELM		0.38	0.26	−0.31	−3.34***
	SARIMA-KF			−0.31	−0.48	−3.81***
	TDNN				−0.36	−3.17***
	SVR					−1.48*

Note: ***, null hypothesis is rejected at the 0.01 level (2-tailed); **, null hypothesis is rejected at the 0.05 level (2-tailed); *, null hypothesis is rejected at the 0.1 level (2-tailed).

Table 7
DM test results for different models for three-step-ahead forecasting.

Vegetable type	Tested model	Reference model				
		ELM	SARIMA-KF	TDNN	SVR	SARIMA
Cabbage	STL-ELM	−1.71**	−1.67**	−1.59*	−1.78**	−5.27 ***
	ELM		0.84	−0.36	−0.31	−3.48***
	SARIMA-KF			0.24	−0.45	−3.84***
	TDNN				−0.38	−3.21 ***
	SVR					−3.05***
Pepper	STL-ELM	−1.87**	−1.38*	−1.57**	−1.75 **	−4.41***
	ELM		0.49	0.34	0.24	−3.53***
	SARIMA-KF			−0.27	−0.42	−3.85***
	TDNN				0.15	−3.38 ***
	SVR					−3.12 ***
Cucumber	STL-ELM	−1.82**	−1.81**	−1.69**	−1.73**	−5.14***
	ELM		0.28	0.38	0.12	−3.41***
	SARIMA-KF			−0.54	−0.31	−3.85***
	TDNN				0.28	−3.62 ***
	SVR					−2.95***
Green bean	STL-ELM	−1.65**	−1.72**	−1.71**	−1.52 *	−5.24***
	ELM		−0.51	−0.47	0.13	−4.68***
	SARIMA-KF			0.24	−0.47	−4.37***
	TDNN				−0.34	−4.82***
	SVR					−4.24***
Tomato	STL-ELM	−1.54*	−1.41*	−1.57**	−1.39 **	−4.87***
	ELM		0.54	0.59	0.47	−4.08***
	SARIMA-KF			−0.21	−0.29	−4.15***
	TDNN				0.24	−4.25***
	SVR					−3.87***

Note: ***, null hypothesis is rejected at the 0.01 level (2-tailed); **, null hypothesis is rejected at the 0.05 level (2-tailed); *, null hypothesis is rejected at the 0.1 level (2-tailed).

Table 8
DM test results for different models for six-step-ahead forecasting.

Vegetable type	Tested model	Reference model				
		ELM	SARIMA-KF	TDNN	SVR	SARIMA
Cabbage	STL-ELM	−1.11	−1.20	−1.57**	−1.65**	−3.21***
	ELM		−0.25	−1.63**	−1.69**	−2.97 ***
	SARIMA-KF			−0.34	−0.47	−2.71***
	TDNN				−0.58	−1.95***
	SVR					−2.15***
Pepper	STL-ELM	−1.87 **	−1.43*	−1.68**	−1.93**	−4.37***
	ELM		0.65	0.37	0.12	−3.35***
	SARIMA-KF			−0.31	−0.38	−3.51***
	TDNN				−0.24	−3.52***
	SVR					−3.24***
Cucumber	STL-ELM	−2.14**	−1.48*	−1.95**	−2.16**	−3.57***
	ELM		0.84	0.46	0.12	−3.28***
	SARIMA-KF			0.54	0.68	−3.34***
	TDNN				0.51	−3.52***
	SVR					−3.14***
Green bean	STL-ELM	0.24	−0.53	−1.15*	−1.58**	−3.96***
	ELM		−0.62	−0.54	0.97	−3.59***
	SARIMA-KF			−0.21	−0.41	−3.49***
	TDNN				−0.61	−4.11***
	SVR					−3.85***
Tomato	STL-ELM	−1.82**	−1.75**	−1.71**	−1.95**	−4.25***
	ELM		0.57	0.49	0.51	−3.21 ***
	SARIMA-KF			−0.35	−0.37	−3.84***
	TDNN				0.47	−3.39***
	SVR					−3.12***

Note: ***, null hypothesis is rejected at the 0.01 level (2-tailed); **, null hypothesis is rejected at the 0.05 level (2-tailed); *, Null hypothesis is rejected at the 0.1 level (2-tailed).

To further evaluate the forecasting performance of the examined forecasting models, the DM statistic is used to test the statistical significance of the listed models. Tables 6–8 report the results of the DM test, where S_{DM} statistics are listed with p-values in brackets for the different models for one-, three-, and six-step-ahead forecasting.

From the results in Tables 6–8, we can deduce the following observations. First, generally speaking, the proposed STL-ELM method outperforms the other listed models at a 5% level of statistical significance in all the cases, though there are some exceptions, where the difference between STL-ELM and SARIMA-KF/ELM/SVR/TDNN is not significant. Second, the SARIMA model is inferior to all of the other listed models at a 10% level of statistical significance, with few exceptions. Third, the difference between the SARIMA-KF and ELM is not significant. Fourth, in the three AI models, ELM is not significantly superior to TDNN and SVR, with one exception at a 5% level of statistical significance.

Overall, from the above analysis of the experimental results obtained in this study, several important conclusions can be drawn as follows. (1) The proposed STL-ELM method is generally the best-performing model, relative to the other competitors listed in this study, for vegetable price forecasting in terms of the SAMPE and MASE criteria. The results of the DM test further illustrate that the proposed STL-ELM method outperforms other competitors at a 5% level of statistical significance. (2) In terms of the comparison between the STL-ELM and ELM, the STL-ELM is generally the winner. This indicates that seasonal adjustment of the time series using certain decomposition techniques before further forecasting can effectively improve the forecasting performance in the case of vegetable prices. (3) Due to the highly seasonal and nonlinear pattern hidden in the vegetable price series, nonlinear models with seasonal adjustment are more suitable for forecasting time series with seasonal volatility than linear models. (4) The proposed STL-ELM method can be used as a promising solution for seasonal time series forecasting.

6. Conclusions

Due to the distinct seasonal characteristics of the vegetable prices, this study proposes a novel hybrid model combining STL and ELM for seasonal forecasting of monthly vegetable prices in the Chinese agriculture market. The experimental results obtained show that the proposed STL-ELM method achieves the best prediction performance relative to the listed competitors for short-, medium-, and long-term forecasting. Thus, the proposed STL-ELM method can be used as a promising tool for forecasting vegetable prices with high seasonality.

In addition to vegetable prices, the proposed STL-ELM method might be applied for other difficult seasonal forecasting tasks in the agriculture market, such as fruit and pork price forecasting. Furthermore, this study restricts attention exclusively to point forecasting; the interval forecasting of vegetable prices is of greater value to practitioners in the agriculture market, which requires further study.

Acknowledgments

This work was supported by the Natural Science Foundation of China (Project No. 71501079, 71771101 and 71571080), the China Postdoctoral Science Foundation (Project No. 2015M570648), the Fundamental Research Funds for the Central Universities(Project No.2662014BQ045 and 2662015PY026), the Major Program of National Social Science Foundation of China (Project No.12&ZD048), and the funding from the Modern Agro-industry Technology Research System (Project No. CARS-23-F01).

Appendix

The formulation of SARIMA

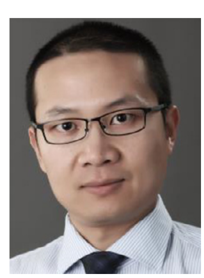
The seasonal ARIMA, proposed by Box and Jenkins [2], has been the most popular method for prediction of time series with strong seasonal pattern. In SARIMA $(p, d, q)\times(P, D, Q)_S$ model, the forecasts are assumed to be a linear combination of past observations and errors. A vegetable price series X_t is SARIMA process with seasonal period length S if d and D are nonnegative integers and if the difference series $W_t(1-B)^d(1-B^S)^D X_t$ is an ARIMA process. The formulation of SARIMA is detailed as follows.

$$\phi_p(B)\Phi_P\left(B^S\right)W_t = \theta_q(B)\Theta_Q\left(B^S\right)\varepsilon_t, t = 1, 2, \ldots, N \tag{9}$$

where N is the number of observations, B is the backshift operator defined by $B^a W_t = W_{t-a}$, $\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ is a non-seasonal autoregressive operator of order p, $\Phi_P(B^S) = 1 - \Phi_1 B^S - \cdots - \Phi_p B^{PS}$ is a seasonal autoregressive operate of order p, $\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ is a non-seasonal moving average operator of order q, $\Theta_Q(B^S) = 1 - \Theta_1 B^S - \cdots - \Theta_Q B^{QS}$ is a seasonal moving average operator of order Q.

References

[1] G. Martin-Rodriguez, J. Caceres-Hernandez, Canary tomato export prices: comparison and relationships between daily seasonal patterns, Span. J. Agric. Res. 11 (2013) 882–893.
[2] G.E. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden–Day series in time series analysis, Holden-Day, 1976.
[3] C. Goh, R. Law, Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention, Tourism Manage. 23 (2002) 499–510.
[4] A. Hayat, M.I. Bhatti, Masking of volatility by seasonal adjustment methods, Econ. Model. 33 (2013) 676–688.
[5] G.P. Zhang, M. Qi, Neural network forecasting for seasonal and trend time series, Eur. J. Oper. Res. 160 (2005) 501–514.
[6] G.P. Zhang, D.M. Kline, Quarterly time-series forecasting with neural networks, IEEE Trans. Neural Netw. 18 (2007) 1800–1814.
[7] W. Hong, Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm, Neurocomputing 74 (2011) 2096–2107.
[8] P. Pai, K. Lin, C. Lin, P. Chang, Time series forecasting by a seasonal support vector regression model, Expert Syst. Appl. 37 (2010) 4261–4265.
[9] G. Huang, Q. Zhu, C. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.
[10] C. Vong, W. Ip, P. Wong, C. Chiu, Predicting minority class for suspended particulate matters level by extreme learning machine, Neurocomputing 128 (2014) 136–144.
[11] N.A. Shrivastava, B.K. Panigrahi, A hybrid wavelet-ELM based short term price forecasting for electricity markets, Int. J. Electr. Power Energy Syst. 55 (2014) 41–50.
[12] A. Mozaffari, M. Vajedi, N.L. Azad, A robust safety-oriented autonomous cruise control scheme for electric vehicles based on model predictive control and online sequential extreme learning machine with a hyper-level fault tolerance-based supervisor, Neurocomputing 151 (2015) 845–856.
[13] P.K. Wong, Z. Yang, C.M. Vong, J. Zhong, Real-time fault diagnosis for gas turbine generator systems using extreme learning machine, Neurocomputing 128 (2014) 249–257.
[14] T. Choi, C. Hui, N. Liu, S. Ng, Y. Yu, Fast fashion sales forecasting with limited data and time, Decis. Support Syst. 59 (2014) 84–92.
[15] C. Lu, Sales forecasting of computer products based on variable selection scheme and support vector regression, Neurocomputing 128 (2014) 491–499.
[16] M. Termenon, M. Graña, A. Savio, A. Akusok, Y. Miche, K.-M. Björk, et al., Brain MRI morphological patterns extraction tool based on Extreme Learning Machine and majority vote classification, Neurocomputing 174 (2016) 344–351.
[17] H. Rong, J. Wei, J. Bai, G. Zhao, Y. Liang, Adaptive neural control for a class of MIMO nonlinear systems with extreme learning machine, Neurocomputing 149 (2015) 405–414.
[18] R.B. Cleveland, W.S. Cleveland, J.E. McRae, I. Terpenning, STL: a seasonal-trend decomposition procedure based on loess, J. Off. Stat. 6 (1990) 3–73.
[19] M. Theodosiou, Forecasting monthly and quarterly time series using STL decomposition, Int. J. Forecast. 27 (2011) 1178–1195.
[20] G.K. Jha, K. Sinha, Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in India, Neural Comput. Appl. 24 (2014) 563–571.
[21] O.B. Shukur, M.H. Lee, Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA, Renewable Energy 76 (2015) 637–647.
[22] G. Chevillon, Direct multi-step estimation and forecasting, J. Econ. Surv. 21 (2007) 746–785.
[23] A. Dieng, Alternative forecasting techniques for vegetable prices in Senegal, Revue sénégalais de recherches agricoles et agroallementalress 1 (2008) 5–10.
[24] H. Adanacioglu, M. Yercan, An analysis of tomato prices at wholesale level in Turkey: an application of SARIMA model, Custos egronegócio on line 8 (2012) 52–75.
[25] G. Nasira, N. Hemageetha, Vegetable price prediction using data mining classification technique. Pattern Recognition, in: 2012 International Conference on: IEEE Informatics and Medical Engineering (PRIME), 2012, pp. 99–102.
[26] A. Amiri, M. Bakhshoodeh, B. Najafi, Forecasting Seasonality in Prices of Potatoes and Onions: Challenge between Geostatistical Models, Neuro Fuzzy Approach and Winter Method, 2011 MPRA paper.
[27] P. Melin, A. Mancilla, M. Lopez, W. Trujillo, J. Cota, S. Gonzalez, Modular neural networks with fuzzy integration applied for time series forecasting, in: Analysis and Design of Intelligent Systems Using Soft Computing Techniques, Springer, 2007, pp. 217–225.
[28] N. Hemageetha, G. Nasira, Radial basis function model for vegetable price prediction, in: 2013 International Conference on: IEEE Pattern Recognition Informatics and Mobile Engineering (PRIME), 2013, pp. 424–428.
[29] E.V. Colino, S.H. Irwin, P. Garcia, Improving the accuracy of outlook price forecasts, Agric. Econ. 42 (2011) 357–371.
[30] F. Davenport, C. Funk, Using time series structural characteristics to analyze grain prices in food insecure countries, Food Secur. 7 (2015) 1055–1070.
[31] C. Luo, Q. Wei, L. Zhou, J. Zhang, S. Sun, Prediction of vegetable price based on neural network and genetic algorithm, in: Computer and Computing Technologies in Agriculture IV, Springer, 2010, pp. 672–681.
[32] Z. Li, L. Cui, S. Xu, L. Weng, X. Dong, G. Li, et al., Prediction model of weekly retail price for eggs based on chaotic neural network, J. Integr. Agric. 12 (2013) 2292–2299.
[33] A. Jumah, R.M. Kunst, Seasonal prediction of European cereal prices: good forecasts using bad models? J. Forecast. 27 (2008) 391–406.
[34] H. Kömm, U. Küsters, Forecasting zero-inflated price changes with a Markov switching mixture model for autoregressive and heteroscedastic time series, Int. J. Forecast. 31 (2015) 598–608.
[35] I.A. Onour, B.S. Sergi, Modeling and forecasting volatility in the global food commodity prices (Modelování a Prognózování Volatility Globálních cen Potravinářských Komodit), Agric. Econ.-Czech 57 (2011) 132–139.
[36] C.O. Ribeiro, S.M. Oliveira, A hybrid commodity price-forecasting model applied to the sugar–alcohol sector, Aust. J. Agric. Resour. Econ. 55 (2011) 180–198.
[37] O.A. Ramirez, M. Fadiga, Forecasting agricultural commodity prices with asymmetric-error GARCH models, J. Agric. Resour. Econ. 28 (2003) 71–85.
[38] M. Shih, B. Huang, N. Chiu, C. Chiu, W. Hu, Farm price prediction using case-based reasoning approach—a case of broiler industry in Taiwan, Comput. Electron. Agric. 66 (2009) 70–75.
[39] A.M. Ticlavilca, Multivariate Bayesian Machine Learning Regression for Operation and Management of Multiple Reservoir, Irrigation Canal, and River Systems. (All Graduate Theses and Dissertations) (2010) p. 600.
[40] K.D. West, K. Wong, A factor model for co-movements of commodity prices, J. Int. Money Finance 42 (2014) 289–309.
[41] X. Zhang, T. Hu, R. Brain, Z. Fu, A forecasting support system for aquatic products price in China, Expert Syst. Appl. 28 (2005) 119–126.
[42] H. Zou, G. Xia, F. Yang, H. Wang, An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting, Neurocomputing 70 (2007) 2913–2923.
[43] T. Xiong, Y. Bao, Z. Hu, Beyond one-step-ahead forecasting: evaluation of alternative multi-step-ahead forecasting models for crude oil prices, Energy Econ. 40 (2013) 405–415.
[44] T. Xiong, Y. Bao, Z. Hu, Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting, Knowl.-Based Syst. 55 (2014) 87–100.
[45] Y. Bao, T. Xiong, Z. Hu, PSO-MISMO modeling strategy for multistep-ahead time series prediction, IEEE Trans. Cybern. 44 (2014) 655–668.
[46] F.X. Diebold, R.S. Mariano, Comparing predictive accuracy, J. Bus. Econ. Stat. 13 (1995) 253–263.
[47] I. Bergmann, G. Ramillien, F. Frappart, Climate-driven interannual ice mass evolution in Greenland, Global Planetary Change 82 (2012) 1–11.
[48] R.E. Kalman, A new approach to linear filtering and prediction problems, J. Basic Eng. 82 (1960) 35–45.
[49] Z. Su, J. Wang, H. Lu, G. Zhao, A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting, Energy Convers. Manage. 85 (2014) 443–452.



Tao Xiong is an associate professor at the Department of Agricultural Economics and Management, College of Economics and Management, Huazhong Agricultural University, PR China. He received his Bsc., Msc., and Ph.D. in Management Science and Engineering from Huazhong University of Science and Technology, PR China, in 2008, 2010 and 2014 respectively. His research interests are agricultural commodity spots and future markets analysis.

Chongguang Li is a professor at the Department of Marketing, College of Economics and Management, Huazhong Agricultural University, PR China. He received his Ph.D. in Agricultural Economics and Management from Huazhong Agricultural University, PR China, in 1998. His research interests are vegetable industry system, vegetable price analysis, and agricultural policy.

Yukun Bao is a professor at School of Management, Huazhong University of Science and Technology, PR China. He received his Bsc., Msc., and Ph.D. in Management Science and Engineering from Huazhong University of Science and Technology, PR China, in 1996,1999 and 2002 respectively. His research interests are time series modeling and forecasting, business intelligence and data mining.