Advances in Computational Intelligence (2022) 2:32
https://doi.org/10.1007/s43674-022-00045-9

ORIGINAL ARTICLE

Canola and soybean oil price forecasts via neural networks

Xiaojie Xu[1] · Yun Zhang[1]

Received: 27 April 2022 / Revised: 24 July 2022 / Accepted: 1 September 2022 / Published online: 15 September 2022

Abstract
Forecasts of commodity prices are vital issues to market participants and policy-makers. Those of cooking section oil are of no exception, considering its importance as one of main food resources. In the present study, we assess the forecast problem using weekly wholesale price indices of canola and soybean oil in China during January 1, 2010–January 3, 2020, by employing the non-linear auto-regressive neural network as the forecast tool. We evaluate forecast performance of different model settings over algorithms, delays, hidden neurons, and data splitting ratios in arriving at the final models for the two commodities, which are relatively simple and lead to accurate and stable results. Particularly, the model for the price index of canola oil generates relative root mean square errors of 2.66, 1.46, and 2.17% for training, validation, and testing, respectively, and the model for the price index of soybean oil generates relative root mean square errors of 2.33, 1.96, and 1.98% for training, validation, and testing, respectively. Through the analysis, we show usefulness of the neural network technique for commodity price forecasts. Our results might serve as technical forecasts on a standalone basis or be combined with other fundamental forecasts for perspectives of price trends and corresponding policy analysis.

Keywords Canola oil · Soybean oil · Price forecast · Time series data · Neural network technique

1 Introduction

Forecasts of commodity prices are vital issues to market participants, which include speculators, processors, hedgers, the media, economists, and policy-makers (Xu 2017a, 2018e). Those of cooking oil prices are of no exception, considering its importance as one of main food resources (Yaakob et al. 2013; Xu and Zhang 2022h; Lam et al. 2010). Due to price volatilities that are generally irregular (Xu 2017c, 2020; Minot 2014; Piot-Lepetit and M'Barek 2011), influences on decisioning processes with great magnitude (Xu and Thurman 2015b; Zhang et al. 2012; Xu 2014c; Mathios 1998; Wells and Slade 2021; Xu and Zhang 2022k), and hence on allocations of resources and economic welfare (Zhang et al. 2014; Xu 2019a, b; Yitzhaki and Slemrod 1991; Ajanovic 2011), significance of forecasting cooking oil prices to the society might not need too much motivation (Caldeira et al. 2019; Xu and Thurman 2015a; Yu et al. 2006; Li et al. 2020b; Brookes et al. 2010).

Researchers in econometrics have devoted significant amounts of efforts to accurate and stable commodity price forecasts. To achieve this goal, a large number of previous studies (Kling and Bessler 1985; Bessler 1982; Brandt and Bessler 1981, 1982, 1983, 1984; Bessler and Chamberlain 1988; Xu and Zhang 2022i; McIntosh and Bessler 1988; Bessler and Brandt 1981; Bessler 1990; Bessler and Babula 1987; Xu 2014b, 2015a; Yang et al. 2001; Bessler et al. 2003; Bessler and Brandt 1992; Bessler and Hopkins 1986; Chen and Bessler 1987, 1990; Wang and Bessler 2004; Bessler and Kling 1986; Babula et al. 2004; Yang et al. 2003; Awokuse and Yang 2003; Yang and Awokuse 2003; Yang and Leatham 1998; Yang et al. 2021) have explored various types of (time series) econometric models and predictions from experts and commercial services. Common time series models in the literature for this forecast purpose include the auto-regressive integrated moving average model (ARIMA), vector auto-regressive model (VAR), vector error correction model (VECM), and different types of their variations. For example, the ARIMA has shown its immense popularity in earlier work and is still being actively sought for many different kinds of time series forecast tasks. It was found that the ARIMA substantially outperforms forecasts based upon expert opinions and structural models for U.S. hog and

✉ Xiaojie Xu
xxu6@ncsu.edu

Yun Zhang
yzhang43@ncsu.edu

[1] North Carolina State University, Raleigh, NC 27695, USA

cattle prices (Brandt and Bessler 1981, 1983; Bessler and Brandt 1981). Further research (Brandt and Bessler 1982, 1984; Kling and Bessler 1985; Bessler 1990) determined that there exists limited space for accuracy improvements for hog price forecasts when changing from the ARIMA to models incorporating more information from the sows farrowing price. This empirical evidence is somewhat different for wheat, for whose prices it was determined that more information from the exchange rate series can benefit improving forecast accuracy obtained via the ARIMA (Bessler and Babula 1987). For canola prices, the ARIMA was also found to achieve decent forecasts (Sulewski et al. 1994). Rather than using one single information source, previous work also suggested the potential value to forecast accuracy by combining the ARIMA with other model types (Bessler and Chamberlain 1988; McIntosh and Bessler 1988). The VAR represents another important econometric method for forecasts of price series that builds upon various economic variables' relations (Bessler and Hopkins 1986; Chen and Bessler 1987; Bessler and Brandt 1992; Awokuse and Yang 2003). The VAR was compared with structural models for the forecast problem of U.S. cotton prices and it was found that the former tends to beat the latter during periods with normal price volatilities (Chen and Bessler 1990). It was demonstrated that the VAR can be useful in sorting out the predictive content among a set of wheat futures prices from different countries (Yang et al. 2003) and U.S. soy and soybean prices of different regions (Babula et al. 2004). Closely related to the VAR, the VECM further includes the long-run relationship(s) among economic variables via cointegration and it could be particularly helpful for long-term price forecasts (Yang and Leatham 1998; Yang and Awokuse 2003; Xu 2019a, b; Yang et al. 2021). For example, it was found that the VECM generally beats the VAR for international wheat price forecasts (Bessler et al. 2003). The general benefit of using the VECM instead of some other models was also determined for several different agricultural price series (Wang and Bessler 2004).

Recently, machine learning techniques have revealed their great potential for price and yield forecasts of a wide spectrum of agricultural commodities (Yuan et al. 2020; Rl and Mishra 2021; Bayona-Oré et al. 2021; Storm et al. 2020; Kouadio et al. 2018; Abreham 2019; Huy et al. 2019; Degife and Sinamo 2019; Naveena et al. 2017; Lopes 2018; Mayabi 2019; Moreno and Salazar 2018; Zelingher et al. 2021; Shahhosseini et al. 2021, 2020; dos Reis Filho 2020; Zelingher et al. 2020; Ribeiro et al. 2019; Surjandari et al. 2015; Ayankoya et al. 2016; Ali et al. 2018; Fang et al. 2020; Harris 2017; Li et al. 2020a; Yoosefzadeh-Najafabadi et al. 2021; Ribeiro and dos Santos 2020; Zhao 2021; Jiang et al. 2019; Handoyo and Chen 2020; Silalahi 2013; Li et al. 2020b; Ribeiro and Oliveira 2011; Zhang et al. 2021; Melo et al. 2007; de Melo et al 2004; Kohzadi et al. 1996; Zou et al. 2007; Rasheed et al. 2021; Khamis and Abdullah 2014; Dias and Rocha 2019; Gómez et al. 2021; Silva et al. 2019; Deina et al. 2011; Filippi et al. 2019; Wen et al. 2021), such as soybeans (dos Reis Filho 2020; Li et al. 2020a; Yoosefzadeh-Najafabadi et al. 2021; Ribeiro and dos Santos 2020; Zhao 2021; Jiang et al. 2019; Handoyo and Chen 2020), soybean oil (Silalahi 2013; Li et al. 2020b), sugar (Surjandari et al. 2015; Ribeiro and Oliveira 2011; Zhang et al. 2021; Melo et al. 2007; de Melo et al 2004; Silva et al. 2019), corn (Xu and Zhang 2021f; Mayabi 2019; Moreno and Salazar 2018; Zelingher et al. 2021; Shahhosseini et al. 2021, 2020; dos Reis Filho 2020; Zelingher et al. 2020; Ribeiro et al. 2019; Surjandari et al. 2015; Ayankoya et al. 2016), wheat (Fang et al. 2020; Ribeiro and dos Santos 2020; Kohzadi et al. 1996; Zou et al. 2007; Rasheed et al. 2021; Khamis and Abdullah 2014; Dias and Rocha 2019; Gómez et al. 2021), coffee (Kouadio et al. 2018; Abreham 2019; Huy et al. 2019; Degife and Sinamo 2019; Naveena et al. 2017; Lopes 2018; Deina et al. 2011), oats (Harris 2017), cotton (Ali et al. 2018; Fang et al. 2020), and canola (Shahwan and Odening 2007; Filippi et al. 2019; Wen et al. 2021). The techniques include the neural network (Xu and Zhang 2021f; Yuan et al. 2020; Abreham 2019; Huy et al. 2019; Naveena et al. 2017; Mayabi 2019; Moreno and Salazar 2018; Ayankoya et al. 2016; Fang et al. 2020; Harris 2017; Li et al. 2020a; Yoosefzadeh-Najafabadi et al. 2021; Ribeiro and dos Santos 2020; Silalahi 2013; Li et al. 2020b; Ribeiro and Oliveira 2011; Zhang et al. 2021; Melo et al. 2007; de Melo et al 2004; Kohzadi et al. 1996; Zou et al. 2007; Rasheed et al. 2021; Khamis and Abdullah 2014; Silva et al. 2019; Deina et al. 2011; Shahwan and Odening 2007), genetic programming Ali et al. (2018), extreme learning (Kouadio et al. 2018; Jiang et al. 2019; Silva et al. 2019; Deina et al. 2011), deep learning (Rl and Mishra 2021), K-nearest neighbor (Abreham 2019; Lopes 2018; Gómez et al. 2021), support vector regression (Abreham 2019; Lopes 2018; dos Reis Filho 2020; Surjandari et al. 2015; Fang et al. 2020; Harris 2017; Li et al. 2020a; Yoosefzadeh-Najafabadi et al. 2021; Ribeiro and dos Santos 2020; Zhao 2021; Li et al. 2020b; Zhang et al. 2021; Dias and Rocha 2019; Gómez et al. 2021), random forest (Kouadio et al. 2018; Lopes 2018; Zelingher et al. 2021; Shahhosseini et al. 2021, 2020; Zelingher et al. 2020; Li et al. 2020a; Yoosefzadeh-Najafabadi et al. 2021; Ribeiro and dos Santos 2020; Dias and Rocha 2019; Gómez et al. 2021; Filippi et al. 2019; Wen et al. 2021), multivariate adaptive regression splines (Dias and Rocha 2019), decision tree (Abreham 2019; Degife and Sinamo 2019; Lopes 2018; Zelingher et al. 2021, 2020; Surjandari et al. 2015; Harris 2017; Dias and Rocha 2019), ensemble (Shahhosseini et al. 2021, 2020; Ribeiro et al. 2019; Fang et al. 2020; Ribeiro and dos Santos 2020), and boosting (Lopes 2018; Zelingher et al. 2021; Shahhosseini et al. 2021, 2020; Zelingher et al. 2020; Ribeiro and dos Santos 2020; Gómez et al. 2021). Efforts have been seen in the literature aiming at improving efficiency and performance of machine learning techniques

(Vajda and Santosh 2016; Elliott et al. 2020). For example, a fast method has been proposed to classify patterns when a k-nearest neighbor classifier is used and the method has been found to not only improve efficiency but also maintain classification performance (Vajda and Santosh 2016). In another work (Elliott et al. 2020), an ensemble method has been successfully constructed to improve efficiency and performance of deep Q-learning. For soybean oil, the genetic algorithm was used to optimize the topology in the neural network for forecasting its prices (Silalahi 2013). It was also found that wavelet transformations and exponential smoothing could benefit forecast accuracy of the neural network and support vector regression for soybean oil futures prices (Li et al. 2020b). For canola, the neural network was combined with the ARIMA to improve price forecast accuracy from each individual model (Shahwan and Odening 2007). In terms of canola's yields, the random forest was successfully adopted for their forecasts (Filippi et al. 2019; Wen et al. 2021). Based on our review here and that from Bayona-Oré et al. (2021), the neural network appears to be the most commonly considered machine learning model for agricultural commodity price forecasting. In particular, previous work (Xu 2015b, 2018a, b, c; Yang et al. 2008, 2010; Wang and Yang 2010; Karasu et al. 2020; Wegener et al. 2016) has revealed that neural network techniques have great potential for forecasts of economic and financial time series, which can be rather noised and chaotic. Previous research (Xu and Zhang 2022j; Yang et al. 2008, 2010; Wang and Yang 2010; Wegener et al. 2016; Karasu et al. 2017a, b) has also demonstrated that neural networks could generate high accuracy across various forecasting circumstances. This might benefit from neural networks' capabilities of self-learning for forecasts (Karasu et al. 2020; Xu and Zhang 2022a) and capturing non-linear characteristics (Altan et al. 2021) in economic and financial data (Xu 2018d; Xu and Zhang 2021a, b). The present study will concentrate on the neural network for forecasting price indices of canola and soybean oil.

To facilitate our analysis, we assess the forecast problem using weekly wholesale price indices of canola and soybean oil in China during January 1, 2010–January 3, 2020 by employing the non-linear auto-regressive neural network as the forecast tool. We evaluate forecast performance of different model settings over algorithms, delays, hidden neurons, and data splitting ratios in arriving at the final models for the two commodities, which are relatively simple and lead to accurate and stable results. To our knowledge and based upon the previous work mentioned above, this is the first study on forecasts of these two vital cooking oil price indices in the Chinese market. There should be little double that it is of great importance to investors and policy makers to have a good understanding of timely and accurate forecasts of commodity prices, which could benefit prompt portfolio adjustments, risk monitoring, and market assessments. By examining the forecast problem using the weekly data, the current study helps timely decisioning. Our results might serve as technical forecasts on a standalone basis or be combined with other fundamental forecasts for perspectives of price trends and associated policy analysis. The forecast framework might also have the potential to be generalized to related forecast problems of other agricultural commodities and in other economic sectors, such as the energy, metal, and mineral.

2 Data

Weekly wholesale price indices of canola and soybean oil in China during January 1, 2010–January 3, 2020 for analysis are plotted in the top panel of Fig. 1, together with their first differences. The average weekly price in June 1994 is used as the base period price index, which is set at 100 that measures the price of canola oil or soybean oil of 50 kilograms. The bottom panel of Fig. 1 also visualizes price indices and their first differences with histograms of fifty bins and kernel estimates to present their distributions. Table 1 reports summary statistics of the data that include the minimum, mean, median, maximum, standard deviation (Std), skewness, kurtosis, and p-value of the Jarque–Bera test of the price indices and their first differences, where one could see that they are not normally distributed, as generally expected for financial series (Xu 2017b, 2019c; Xu and Zhang 2022b). It is worth noting that price indices of soybean oil are missing on February 19, 2010, and February 3, 2017, and we utilize cubic spline interpolation for approximations of 99.39 and 102.52, respectively. The approximated price index of 99.39 on February 19, 2010, is close to 101.84 on February 12, 2010, and 101.40 on February 26, 2010. Similarly, the approximated price index of 102.52 on February 3, 2017, is close to 103.49 on January 27, 2017, and 99.76 on February 10, 2017.

3 Methods

We use the non-linear auto-regressive neural network model as the forecasting tool for weekly price indices of canola and soybean oil. This model could be expressed as $y_t = f(y_{t-1}, ..., y_{t-d})$. Here, y is the weekly price index of canola or soybean oil to be forecasted, t denotes time, d denotes the number of delays, and f denotes the function. We concentrate on one-week ahead forecasts.

We adopt the two-layer feedforward network, which has a sigmoid transfer function among hidden layers and a linear transfer function associated with the output layer. In terms of algorithms for model training, we consider both the Levenberg–Marquardt (LM) algorithm (Levenberg 1944; Marquardt 1963) and scaled conjugate gradient (SCG) algorithm (Møller 1993), which have been employed in a diverse

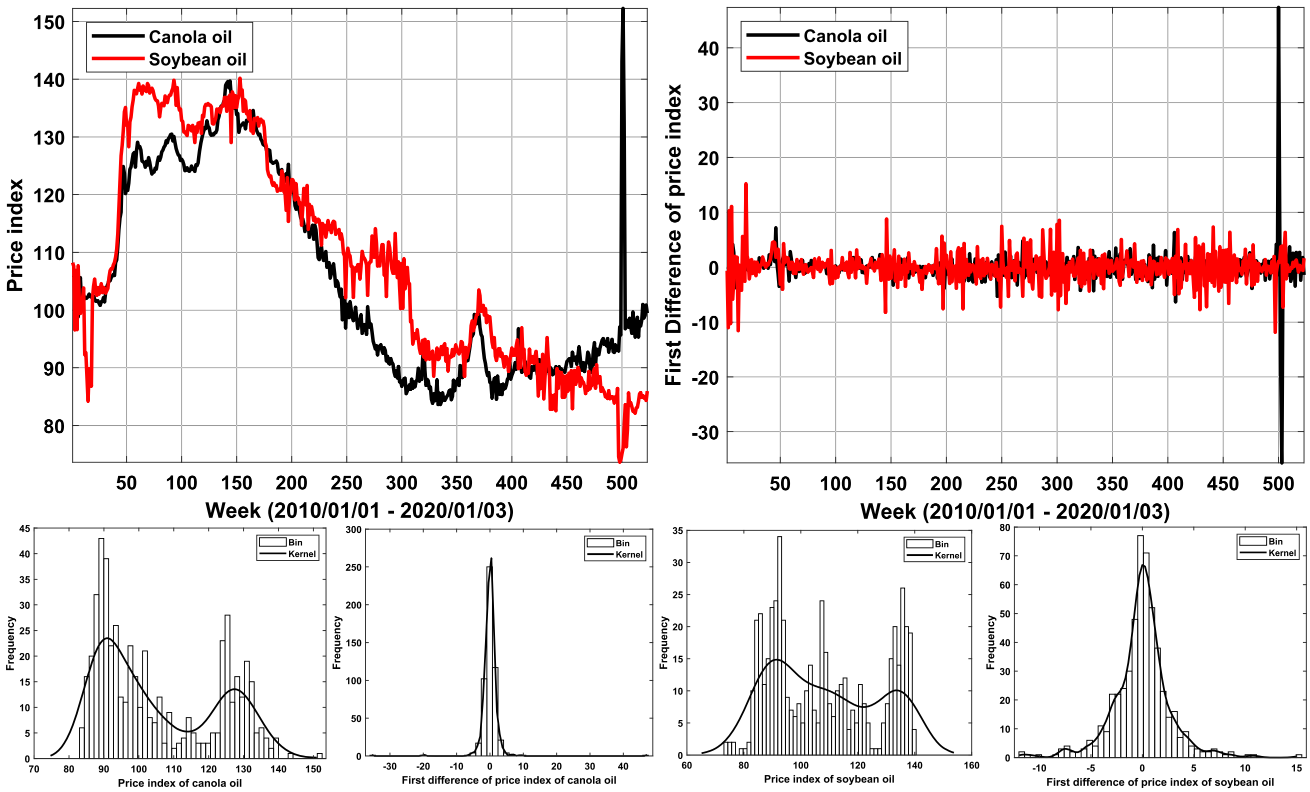


Fig. 1 Top panel: weekly price indices of canola and soybean oil (left) and first differences of prices (right); Bottom panel: histograms of fifty bins and kernel estimates for weekly price indices and their first differences of canola oil (left two) and those of soybean oil (right two)

Table 1 Summary statistics of weekly price indices and their first differences of canola and soybean oil

Commodity	Series	Minimum	Mean	Median	Std	Maximum	Skewness	Kurtosis	Jarque–Bera
Canola oil	Price	83.6200	105.5743	99.2000	16.9969	152.2600	0.5020	1.7533	< 0.001
	First difference	−35.6800	−0.0118	0.0500	3.1627	47.3600	3.2782	131.2897	< 0.001
Soybean oil	Price	73.6600	108.1831	106.3700	18.7667	140.1700	0.2779	1.6938	< 0.001
	First difference	−11.8600	−0.0427	0.0150	2.8143	15.2100	0.0307	7.2074	< 0.001

variety of research fields (Xu and Zhang 2021c, d, 2022c, 2021e; Doan and Liong 2004; Kayri 2016; Khan et al. 2019; Selvamuthu et al. 2019). These two algorithms have been found to be useful for forecasting time series with nonlinear patterns (Abraham 2004; Asadi et al. 2012; Ahadi and Liang 2018; Selvamuthu et al. 2019; Qazani et al. 2021). Comparative studies of these two algorithms could be found from some of previous work (Baghirli 2015; Xu and Zhang 2022d, e; Al Bataineh and Kaur 2018). The LM algorithm makes approximations of the second-order training speed so that it could avoid expensive computing of the Hessian matrix, H (Paluszek and Thomas 2020), and it could efficiently handle the slow convergence issue (Hagan and Menhaj 1994). The SCG algorithm avoids time-consuming line searches in conjugate gradient algorithms and is generally faster as compared to the LM backpropagation.

Specifically, the approximation performed through the LM algorithm could be expressed as $H = J^T J$, where $J = \left[\frac{\partial f}{\partial z_1} \; \frac{\partial f}{\partial z_2} \right]$ for a nonlinear function $f(z_1, z_2)$ with $H = \begin{bmatrix} \frac{\partial^2 f}{\partial z_1^2} & \frac{\partial^2 f}{\partial z_1 \partial z_2} \\ \frac{\partial^2 f}{\partial z_2 \partial z_1} & \frac{\partial^2 f}{\partial z_2^2} \end{bmatrix}$. $g = J^T e$ is used to represent the gradient, where e contains the error vector. The rule of $z_{k+1} = z_k - \left[J^T J + \mu I \right]^{-1} J^T e$ is utilized to make updates of weights and biases, where I represents the identity matrix. The LM algorithm is similar to Newton's method for the case of $\mu = 0$ and it is gradient descent with small step sizes when μ is large. The value of μ will be decreased if faster gradient descent is less needed after successful steps. The LM algorithm not only has desired properties of steepest-descent algorithms and Gauss–Newton methods but also avoids some of their limitations that include the potential issue of slow convergence (Hagan and Menhaj 1994).

Backpropagation algorithms carry out adjustments of weights in the steepest descent as the performance function would rapidly decrease in the direction, which however, does not always reflect the fastest convergence. Conjugate gradient algorithms carry out searches along the conjugate direction, which in general, result in faster convergence as compared to the steepest descent. Most algorithms would apply learning rates for determining lengths of updated weight step sizes. For conjugate gradient algorithms, step sizes are modified during iterations. Hence, the search is carried out along the conjugate gradient direction for determining the step size for the reduction of the performance function. Besides, for the purpose of avoiding time-consuming line searches in conjugate gradient algorithms, the SCG algorithm, which is fully-automated and supervised, could be used. It is generally quicker than the LM backpropagation.

In addition to different algorithms considered, different model settings over delays, hidden neurons, and data spitting ratios are examined as well. We consider delays of 2, 3, 4, 5, and 6, hidden neurons of 2, 3, 5 and 10, and data spitting ratios of 60% vs. 20% vs. 20%, 70% vs. 15% vs. 15%, and 80% vs. 10% vs. 10% for training, validation, and testing. These model settings are summarized in Table 2, where the setting #27 is utilized to build the final chosen model for the price index of canola oil and the setting #23 for the price index of soybean oil, both of which are trained through the LM algorithm following the ratio of 70% vs. 15% vs. 15% for training, validation, and testing. The setting #27 is based on 5 delays and 5 hidden neurons, and the setting #23 is based on 3 delays and 5 hidden neurons.

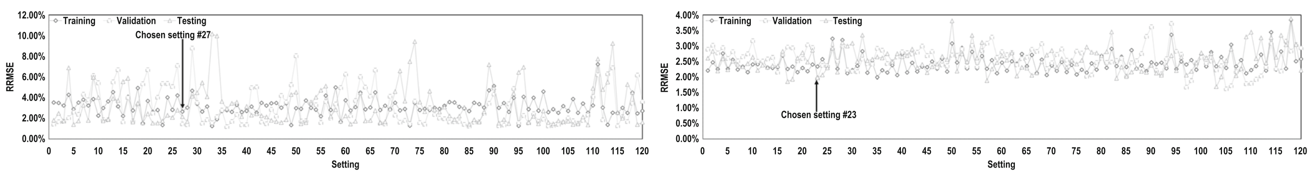
4 Results

All of the model settings shown in Table 2 are run for weekly price indices of canola and soybean oil. For each model setting, the relative root mean square error (RRMSE), as the forecast performance metric, is computed for training, validation, and testing phases and the corresponding results are shown in Fig. 2. The RRMSE expresses forecast performance in a percentage form and helps comparisons of forecast performance across different series and models. Balancing forecast performance and stabilities, the setting #27 (5 delays and 5 hidden neurons) is chosen for the price index of canola oil and the setting #23 (3 delays and 5 hidden neurons) for the price index of soybean oil, both of which are based upon the LM algorithm and the data splitting ratio of 70% vs. 15% vs. 15% for training, validation, and testing. These two chosen settings are marked with dark arrows in Fig. 2 and one should be able to observe that they not only generate rather low RRMSEs but also produce rather close RRMSEs. More specifically, one could see from Fig. 2 that for the chosen settings, the diamond corresponding to training, square corresponding to validation, and triangular corresponding to testing are pretty close to each other. Taking canola oil as an example, there exist other settings with a lower RRMSE as compared to the setting #27 for a specific sub-sample but with higher RRMSEs for the remaining sub-samples, meaning a lower stability. For example, the setting #15 shows a slightly lower RRMSE than the setting #27 for training but higher RRMSEs for validation and testing, as well as a higher overall RRMSE. Choosing the model setting with relatively stable performance across training, validation, and testing could help ensure no overfitting or underfitting.

Having the chosen setting determined for each commodity, performance sensitivities to different settings are evaluated through altering one setting a time and the corresponding results are presented in Fig. 3, where RRMSEs for training, validation, and testing based upon each setting are reported. For the price index of canola oil, the comparison between the settings #27 and #28 evaluates the sensitivity to algorithms, between the setting #27 and settings #21, #23, #25, and #29 the sensitivity to delays, between the setting #27 and settings #7, #17, and #37 the sensitivity to hidden neurons, and between the setting #27 and settings #67 and #107 the sensitivity to data splitting ratios. For the price index of soybean oil, the comparison between the settings #23 and #24 evaluates the sensitivity to algorithms, between the setting #23 and settings #21, #25, #27, and #29 the sensitivity to delays, between the setting #23 and settings #3, #13, and #33 the sensitivity to hidden neurons, and between the setting #23 and settings #63 and #103 the sensitivity to data splitting ratios. These results support the settings #27 as the final choice for the price index of canola oil, leading to RRMSEs of 2.66, 1.46, and 2.17% for training, validation, and testing, respectively, and the overall RRMSE of 2.45%, and these results support the settings #23 as the final choice for the price index of soybean oil, leading to RRMSEs of 2.33, 1.96, and 1.98% for training, validation, and testing, respectively, and the overall RRMSE of 2.23%. From the perspective of the mean absolute error (MAE), the setting #27 leads to MAEs of 1.2512, 1.1289, and 1.3776 for training, validation, and testing, respectively, and the overall MAE of 1.2518 for the price index of canola oil, and the setting #23 leads to MAEs of 1.7565, 1.6153, and 1.5667 for training, validation, and testing, respectively, and the overall MAE of 1.7069 for the price index of soybean oil. We could observe from Fig. 3 that the settings #27 and #23 lead to rather stable performance across the training, validation, and testing phases among the alternatives for the price indices of canola oil and soybean oil, respectively. From Fig. 3, it could be seen that better overall performance is achieved through the LM algorithm as compared to the SCG algorithm, which is reflected through the comparison between the settings #27 that is based on the LM

Table 2 Explored model settings for weekly price indices of canola and soybean oil

		Model setting
Algorithm	LM	$1 + 2i(i = 0, 1, ..., 59)$
	SCG	$2 + 2i(i = 0, 1, ..., 59)$
Delay	2	$1 + 10j - 2 + 10j(j = 0, 1, ..., 11)$
	3	$3 + 10j - 4 + 10j(j = 0, 1, ..., 11)$
	4	$5 + 10j - 6 + 10j(j = 0, 1, ..., 11)$
	5	$7 + 10j - 8 + 10j(j = 0, 1, ..., 11)$
	6	$9 + 10j - 10 + 10j(j = 0, 1, ..., 11)$
Hidden neuron	2	$1 + 40k - 10 + 40k(k = 0, 1, 2)$
	3	$11 + 40k - 20 + 40k(k = 0, 1, 2)$
	5	$21 + 40k - 30 + 40k(k = 0, 1, 2)$
	10	$31 + 40k - 40 + 40k(k = 0, 1, 2)$
Training vs. validation vs. testing ratio	70% vs. 15% vs. 15%	1–40
	60% vs. 20% vs. 20%	41–80
	80% vs. 10% vs. 10%	81–120



(a) Canola oil **(b)** Soybean oil

Fig. 2 RRMSEs across all model settings for weekly price indices of canola and soybean oil

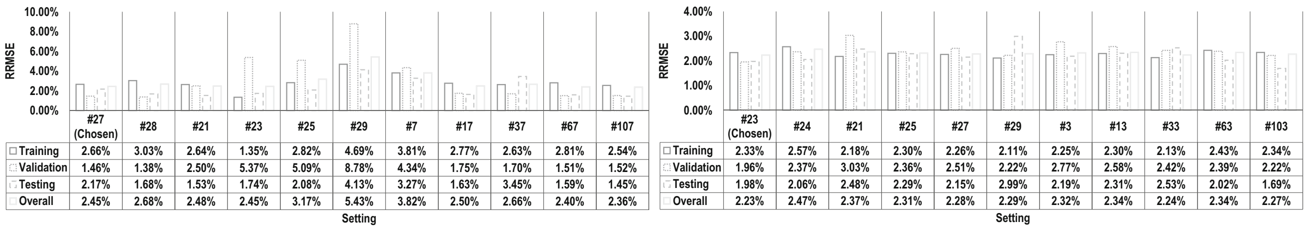
algorithm and #28 that is based on the SCG algorithm for the price index of canola oil and through the comparison between the settings #23 that is based on the LM algorithm and #24 that is based on the SCG algorithm for the price index of soybean oil. This is consistent with the literature (Xu and Zhang 2022l, f; Batra 2014), which finds that while the SCG algorithm is generally better in terms of speed than the LM algorithm on a multilayer perceptron structure with two hidden layers, the LM algorithm generally leads to slightly better performance in terms of accuracy than the SCG algorithm.

Detailed visualization of forecasted results and forecast errors based upon the chosen setting for the training, validation, and testing phases are shown in Fig. 4 for each commodity. Overall, the chosen setting results in accurate and stable performance, suggesting usefulness of the neural network technique for forecasting weekly price indices of canola and soybean oil. Particularly, from Fig. 4 (top panel), we could observe that the forecasted price indices closely track the observed ones across the training, validation, and testing phases. From Fig. 4 (bottom panel), we could see that there is no consistent overprediction or underprediction across the training, validation, and testing phases. One could also observe that a couple of forecast errors shown in Fig. 4 (bottom panel) are larger during periods with significantly elevated price volatilities, particularly for canola oil near the end of the sample. This might not be surprising and the models generally still capture the trends during these periods.

5 Discussion

We have conducted error autocorrelation analysis as well (details available upon request) and autocorrelations associated with different lags up to the lag of 20 are all within the 95% confidence limits except for the lags of 6 and 13 for canola and soybean oil, respectively, for which slight breaches of the confidence limit are found. These slight breaches will be avoided if the 99% confidence limit is used. The error autocorrelation analysis thus suggests that the chosen settings are generally adequate.

It has been well established in the literature (Yang et al. 2008, 2010; Wang and Yang 2010; Karasu et al. 2020) that there could be nonlinearities in higher moments inhabiting financial and economic time series data. We apply the BDS test (Brock et al. 1996), for which one might refer to Dergiades et al. (2013) and Fujihara and Mougoué (1997) for a formal description and to Brock et al. (1996) for all technical details, to the weekly price indices of canola and soybean oil examined in the current study and find that p values of the tests are all well below 0.01 and almost 0 based upon



(a) Canola oil **(b)** Soybean oil

Fig. 3 Sensitivities of model performance (the RRMSE) to different model settings for weekly price indices of canola and soybean oil

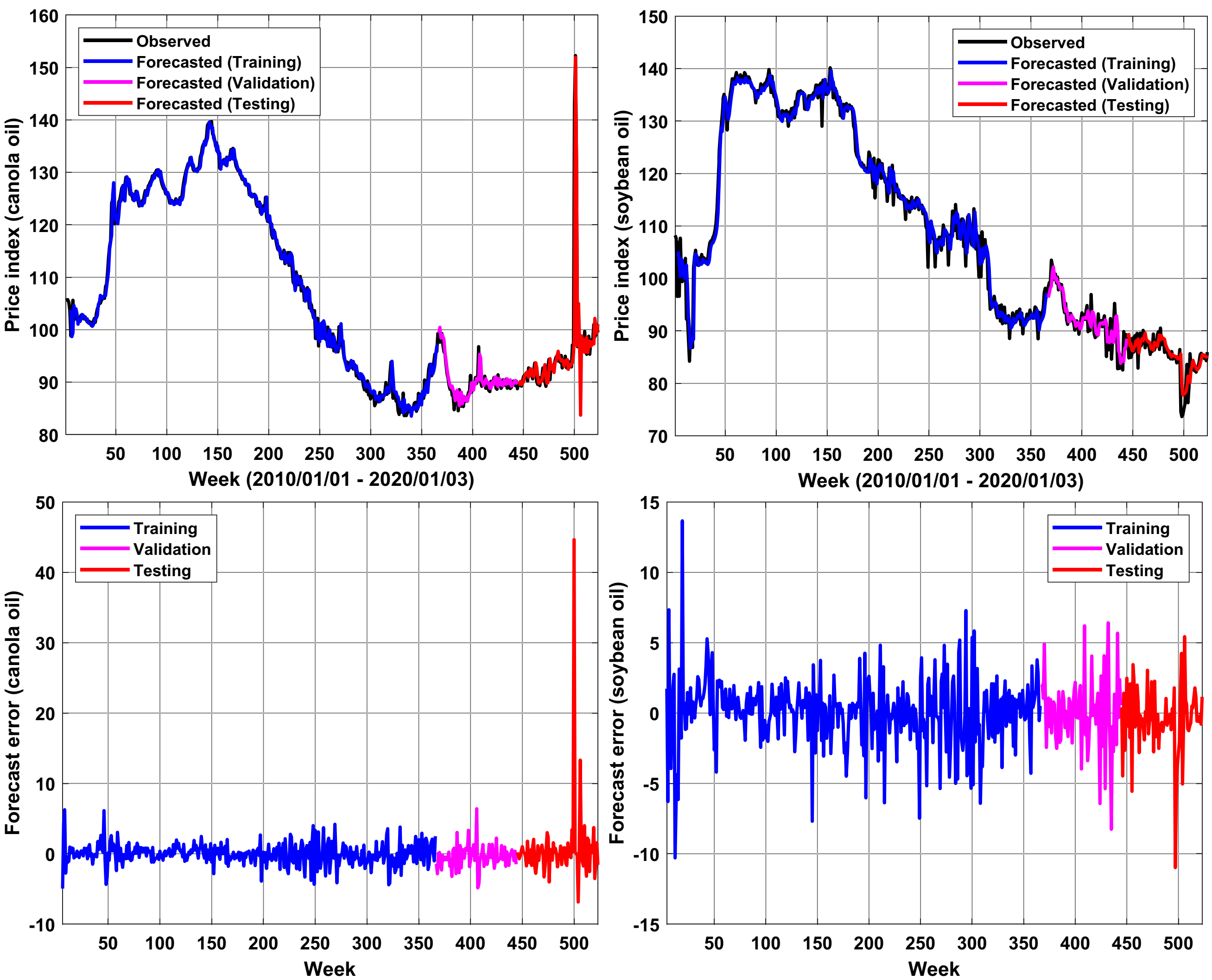


Fig. 4 Top panel: forecasts of weekly price indices of canola and soybean oil; Bottom panel: forecast errors calculated as observations minus forecasts

embedding dimensions of 2 to 10 and ϵ values (i.e. distance used for testing proximity of data points) of 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 times the standard deviation of the price index series. Neural network techniques have capabilities of self-learning for forecasts (Karasu et al. 2020) and capturing non-linear features (Altan et al. 2021) often inhabiting financial and economic series, such as the cooking oil price indices considered here. The neural network's one advantage over other non-linear techniques for time series modeling is that it would well approximate a large class of functions with a class of multi-layer neural networks (Yang et al. 2008, 2010; Wang and Yang 2010). Unlike common non-linear models that employ a specific non-linear function between inputs and the output, the neural network's multi-layer structure would combine many 'basic' non-linear functions. With good forecast performance achieved here, usefulness of the neural network technique is empirically demonstrated to the forecast issue of the weekly price indices of canola and soybean oil.

6 Conclusion

Forecasts of commodity prices represent vital issues to market participants and policy makers. Those of cooking oil are of no exception. In the present study, the forecast problem is investigated for weekly wholesale price indices of canola and soybean oil in China during January 1, 2010–January 3, 2020. The forecast technique adopted here is the non-linear auto-regressive neural network and the final models for the two commodities are built by exploring different model settings. For price indices of both commodities, the final models are constructed based upon the Levenberg–Marquardt algorithm (Levenberg 1944; Marquardt 1963) and a data splitting ratio of 70% vs. 15% vs. 15% for training, validation, and testing. The model for the price index of canola oil uses 5 delays and 5 hidden neurons, and that for the price index of soybean oil uses 3 delays and 5 hidden neurons. The models lead to accurate and stable forecast performance. Particularly, the model for the price index of canola oil generates relative root mean square errors (RRMSEs) of 2.66, 1.46, and 2.17% for training, validation, and testing, respectively, and the overall RRMSE of 2.45%, and the model for the price index of soybean oil generates RRMSEs of 2.33, 1.96, and 1.98% for training, validation, and testing, respectively, and the overall RRMSE of 2.23%. Our results might serve as technical forecasts on a standalone basis or be combined with other fundamental forecasts for perspectives of price trends and associated policy analysis. The framework presented here should not appear to be difficult to implement, which can be an important consideration to decision makers (Brandt and Bessler 1983), and it might also have the potential to be generalized to related forecast problems of other agricultural commodities and in other economic sectors, such as the energy, metal, and mineral. Future research of interest might be examining the potential of combining (non)linear time series techniques and graph theory from machine learning for price forecasts (Kano and Shimizu 2003; Shimizu et al. 2006; Xu and Zhang 2022g; Shimizu and Kano 2008; Shimizu et al. 2011; Xu 2014a; Bessler and Wang 2012). Exploring economic significance of adopting neural network modeling or other machine learning techniques for price forecasts might also be a worthwhile avenue for future research (Yang et al. 2008, 2010; Wang and Yang 2010).

Data Availability Statement The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors did not receive support from any organization for the submitted work. The authors have no relevant financial or non-financial interests to disclose.

References

Abraham A (2004) Meta learning evolutionary artificial neural networks. Neurocomputing 56:1–38. https://doi.org/10.1016/S0925-2312(03)00369-2

Abreham Y (2019) Coffee price prediction using machine-learning techniques. Ph.D. thesis, ASTU

Ahadi A, Liang X (2018) Wind speed time series predicted by neural network. In: 2018 IEEE Canadian conference on electrical and computer engineering (CCECE). IEEE, pp 1–4. https://doi.org/10.1109/CCECE.2018.8447635

Ajanovic A (2011) Biofuels versus food production: does biofuels production increase food prices? Energy 36:2070–2076. https://doi.org/10.1016/j.energy.2010.05.019

Al Bataineh A, Kaur D (2018) A comparative study of different curve fitting algorithms in artificial neural network using housing dataset. In: NAECON 2018-IEEE national aerospace and electronics conference. IEEE, pp 174–178. https://doi.org/10.1109/NAECON.2018.8556738

Ali M, Deo RC, Downs NJ, Maraseni T (2018) Cotton yield prediction with Markov chain monte carlo-based simulation model integrated with genetic programing algorithm: a new hybrid copula-driven approach. Agric For Meteorol 263:428–448. https://doi.org/10.1016/j.agrformet.2018.09.002

Altan A, Karasu S, Zio E (2021) A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. Appl Soft Comput 100:106996. https://doi.org/10.1016/j.asoc.2020.106996

Asadi S, Hadavandi E, Mehmanpazir F, Nakhostin MM (2012) Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction. Knowl-Based Syst 35:245–258. https://doi.org/10.1016/j.knosys.2012.05.003

Awokuse TO, Yang J (2003) The informational role of commodity prices in formulating monetary policy: a reexamination. Econ Lett 79:219–224. https://doi.org/10.1016/S0165-1765(02)00331-2

Ayankoya K, Calitz AP, Greyling JH (2016) Using neural networks for predicting futures contract prices of white maize in South Africa. In: Proceedings of the annual conference of the South African Institute of Computer Scientists and Information Technologists, pp 1–10. https://doi.org/10.1145/2987491.2987508

Babula RA, Bessler DA, Reeder J, Somwaru A (2004) Modeling US soy-based markets with directed acyclic graphs and Bernanke structural var methods: the impacts of high soy meal and soybean prices. J Food Distrib Res 35:29–52. https://doi.org/10.22004/ag.econ.27559

Baghirli O (2015) Comparison of Lavenberg–Marquardt, scaled conjugate gradient and bayesian regularization backpropagation algorithms for multistep ahead wind speed forecasting using multilayer perceptron feedforward neural network. https://www.diva-portal.org/smash/get/diva2:828170/FULLTEXT01.pdf

Batra D (2014) Comparison between Levenberg–Marquardt and scaled conjugate gradient training algorithms for image compression using MLP. Int J Image Process (IJIP) 8:412–422

Bayona-Oré S, Cerna R, Hinojoza ET (2021) Machine learning for price prediction for agricultural products. WSEAS Trans Bus Econ 18:969–977. https://doi.org/10.37394/23207.2021.18.92

Bayona-Oré S, Cerna R, Tirado Hinojoza E (2021) Machine learning for price prediction for agricultural products. https://doi.org/10.37394/23207.2021.18.92

Bessler DA (1982) Adaptive expectations, the exponentially weighted forecast, and optimal statistical predictors. A revisit. Agric Econ Res 34:16–23. https://doi.org/10.22004/ag.econ.148819

Bessler DA (1990) Forecasting multiple time series with little prior information. Am J Agric Econ 72:788–792. https://doi.org/10.2307/1243059

Bessler DA, Babula RA (1987) Forecasting wheat exports: do exchange rates matter? J Bus Econ Stat 5:397–406. https://doi.org/10.2307/1391615

Bessler DA, Brandt JA (1981) Forecasting livestock prices with individual and composite methods. Appl Econ 13:513–522. https://doi.org/10.1080/00036848100000016

Bessler DA, Brandt JA (1992) An analysis of forecasts of livestock prices. J Econ Behav Organ 18:249–263. https://doi.org/10.1016/0167-2681(92)90030-F

Bessler DA, Chamberlain PJ (1988) Composite forecasting with Dirichlet priors. Decis Sci 19:771–781. https://doi.org/10.1111/j.1540-5915.1988.tb00302.x

Bessler DA, Hopkins JC (1986) Forecasting an agricultural system with random walk priors. Agric Syst 21:59–67. https://doi.org/10.1016/0308-521X(86)90029-6

Bessler DA, Kling JL (1986) Forecasting vector autoregressions with Bayesian priors. Am J Agric Econ 68:144–151. https://doi.org/10.2307/1241659

Bessler DA, Wang Z (2012) D-separation, forecasting, and economic science: a conjecture. Theor Decis 73:295–314. https://doi.org/10.1007/s11238-012-9305-8

Bessler DA, Yang J, Wongcharupan M (2003) Price dynamics in the international wheat market: modeling with error correction and directed acyclic graphs. J Reg Sci 43:1–33. https://doi.org/10.1111/1467-9787.00287

Brandt JA, Bessler DA (1981) Composite forecasting: an application with US hog prices. Am J Agr Econ 63:135–140. https://doi.org/10.2307/1239819

Brandt JA, Bessler DA (1982) Forecasting with a dynamic regression model: a heuristic approach, North Central. J Agric Econ 4:27–33. https://doi.org/10.2307/1349096

Brandt JA, Bessler DA (1983) Price forecasting and evaluation: an application in agriculture. J Forecast 2:237–248. https://doi.org/10.1002/for.3980020306

Brandt JA, Bessler DA (1984) Forecasting with vector autoregressions versus a univariate Arima process: an empirical example with US hog prices, North Central. J Agric Econ 4:29–36. https://doi.org/10.2307/1349248

Brock WA, Scheinkman JA, Dechert WD, LeBaron B (1996) A test for independence based on the correlation dimension. Economet Rev 15:197–235. https://doi.org/10.1080/07474939608800353

Brookes G, Yu T-H, Tokgoz S, Elobeid A (2010) The production and price impact of biotech corn, canola, and soybean crops

Caldeira C, Swei O, Freire F, Dias LC, Olivetti EA, Kirchain R (2019) Planning strategies to address operational and price uncertainty in biodiesel production. Appl Energy 238:1573–1581. https://doi.org/10.1016/j.apenergy.2019.01.195

Chen DT, Bessler DA (1987) Forecasting the US cotton industry: structural and time series approaches. In: Proceedings of the NCR-134 conference on applied commodity price analysis. Forecasting, and Market Risk Management, Chicago Mercantile Exchange, Chicago. https://doi.org/10.22004/ag.econ.285463

Chen DT, Bessler DA (1990) Forecasting monthly cotton price: structural and time series approaches. Int J Forecast 6:103–113. https://doi.org/10.1016/0169-2070(90)90101-G

de Melo B, Júnior CN, Milioni AZ (2004) Daily sugar price forecasting using the mixture of local expert models. WIT Trans Inf Commun Technol. https://doi.org/10.2495/DATA040221

Degife WA, Sinamo A (2019) Efficient predictive model for determining critical factors affecting commodity price: the case of coffee in Ethiopian Commodity Exchange (ECX). Int J Inf Eng Electron Bus 11:32–36. https://doi.org/10.5815/ijieeb.2019.06.05

Deina C, do Amaral Prates MH, Alves CHR, Martins MSR, Trojan F, Stevan Jr SL, Siqueira HV, (2011) A methodology for coffee price forecasting based on extreme learning machines. Inf Process Agric. https://doi.org/10.1016/j.inpa.2021.07.003

Dergiades T, Martinopoulos G, Tsoulfidis L (2013) Energy consumption and economic growth: parametric and non-parametric causality testing for the case of Greece. Energy Econ 36:686–697. https://doi.org/10.1016/j.eneco.2012.11.017

Dias J, Rocha H (2019) Forecasting wheat prices based on past behavior: comparison of different modelling approaches. In: International conference on computational science and its applications. Springer, pp 167–182. https://doi.org/10.1007/978-3-030-24302-9_13

Doan CD, Liong SY (2004) Generalization for multilayer neural network bayesian regularization or early stopping. In: Proceedings of Asia Pacific Association of hydrology and water resources 2nd conference, pp 5–8

dos Reis Filho IJ, Correa GB, Freire GM, Rezende SO (2020) Forecasting future corn and soybean prices: an analysis of the use of textual information to enrich time-series. In: Anais do VIII symposium on knowledge discovery, mining and learning. SBC, pp 113–120

Elliott DL, Santosh KC, Anderson C (2020) Gradient boosting in crowd ensembles for q-learning using weight sharing. Int J Mach Learn Cybern 11:2275–2287. https://doi.org/10.1007/s13042-020-01115-5

Fang Y, Guan B, Wu S, Heravi S (2020) Optimal forecast combination based on ensemble empirical mode decomposition for agricultural commodity futures prices. J Forecast 39:877–886. https://doi.org/10.1002/for.2665

Filippi P, Jones EJ, Wimalathunge NS, Somarathna PD, Pozza LE, Ugbaje SU, Jephcott TG, Paterson SE, Whelan BM, Bishop TF (2019) An approach to forecast grain crop yield using multilayered, multi-farm data sets and machine learning. Precis Agric 20:1015–1029. https://doi.org/10.1007/s11119-018-09628-4

Fujihara RA, Mougoué M (1997) An examination of linear and nonlinear causal relationships between price variability and volume in petroleum futures markets. J Futures Mark Futures Options Other Deriv Prod 17:385–416. https://doi.org/10.1002/(SICI)1096-9934(199706)17:4385::AID-FUT23.0.CO;2-D

Gómez D, Salvador P, Sanz J, Casanova JL (2021) Modelling wheat yield with antecedent information, satellite and climate data using machine learning methods in Mexico. Agric For Meteorol 300:108317. https://doi.org/10.1016/j.agrformet.2020.108317

Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 5:989–993. https://doi.org/10.1109/72.329697

Handoyo S, Chen YP (2020) The developing of fuzzy system for multiple time series forecasting with generated rule bases and optimized consequence part. SSRG Int J Eng Trends Technol 68:18–122. https://doi.org/10.14445/22315381/IJETT-V68I12P220

Harris JJ (2017) A machine learning approach to forecasting consumer food prices

Huy HT, Thac HN, Thu HNT, Nhat AN, Ngoc VH (2019) Econometric combined with neural network for coffee price forecasting. J Appl Econ Sci 14

Jiang F, He J, Zeng Z (2019) Pigeon-inspired optimization and extreme learning machine via wavelet packet analysis for predicting bulk commodity futures prices. Sci China Inf Sci 62:1–19. https://doi.org/10.1007/s11432-018-9714-5

Kano Y, Shimizu S et al (2003) Causal inference using nonnormality. In: Proceedings of the international symposium on science of modeling, the 30th anniversary of the information criterion, pp 261–270. http://www.ar.sanken.osaka-u.ac.jp/~sshimizu/papers/aic30_web2.pdf

Karasu S, Altan A, Saraç Z, Hacioğlu R (2017a) Prediction of wind speed with non-linear autoregressive (NAR) neural networks. In: 2017 25th signal processing and communications applications conference (SIU). IEEE, pp 1–4. https://doi.org/10.1109/SIU.2017.7960507

Karasu S, Altan A, Saraç Z, Hacioğlu R (2017b) Estimation of fast varied wind speed based on NARX neural network by using curve fitting. Int J Energy Appl Technol 4:137–146. https://dergipark.org.tr/en/download/article-file/354536

Karasu S, Altan A, Bekiros S, Ahmad W (2020) A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series. Energy 212:118750. https://doi.org/10.1016/j.energy.2020.118750

Kayri M (2016) Predictive abilities of Bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data. Math Comput Appl 21:20. https://doi.org/10.3390/mca21020020

Khamis A, Abdullah S (2014) Forecasting wheat price using backpropagation and NARX neural network. Int J Eng Sci 3:19–26

Khan TA, Alam M, Shahid Z, Mazliham M (2019) Comparative performance analysis of Levenberg–Marquardt, Bayesian regularization and scaled conjugate gradient for the prediction of flash floods. J Inf Commun Technol Robot Appl 10:52–58. http://jictra.com.pk/index.php/jictra/article/view/188/112

Kling JL, Bessler DA (1985) A comparison of multivariate forecasting procedures for economic time series. Int J Forecast 1:5–24. https://doi.org/10.1016/S0169-2070(85)80067-4

Kohzadi N, Boyd MS, Kermanshahi B, Kaastra I (1996) A comparison of artificial neural network and time series models for forecasting commodity prices. Neurocomputing 10:169–181. https://doi.org/10.1016/0925-2312(95)00020-8

Kouadio L, Deo RC, Byrareddy V, Adamowski JF, Mushtaq S et al (2018) Artificial intelligence approach for the prediction of Robusta Coffee Yield using soil fertility properties. Comput Electron Agric 155:324–338. https://doi.org/10.1016/j.compag.2018.10.014

Lam MK, Lee KT, Mohamed AR (2010) Homogeneous, heterogeneous and enzymatic catalysis for transesterification of high free fatty acid oil (waste cooking oil) to biodiesel: a review. Biotechnol Adv 28:500–518. https://doi.org/10.1016/j.biotechadv.2010.03.002

Levenberg K (1944) A method for the solution of certain non-linear problems in least squares. Q Appl Math 2:164–168. https://doi.org/10.1090/qam/10666

Li J, Li G, Liu M, Zhu X, Wei L (2020a) A novel text-based framework for forecasting agricultural futures using massive online news headlines. Int J Forecast. https://doi.org/10.1016/j.ijforecast.2020.02.002

Li G, Chen W, Li D, Wang D, Xu S (2020b) Comparative study of short-term forecasting methods for soybean oil futures based on LSTM, SVR, ES and wavelet transformation. In: Journal of physics: conference series, volume 1682. IOP Publishing, p 012007. https://doi.org/10.1088/1742-6596/1682/1/012007

Lopes LP (2018) Prediction of the Brazilian natural coffee price through statistical machine learning models. SIGMAE 7:1–16

Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 11:431–441. https://doi.org/10.1137/0111030

Mathios AD (1998) The importance of nutrition labeling and health claim regulation on product choice: an analysis of the cooking oils market. Agric Resour Econ Rev 27:159–168. https://doi.org/10.1017/S1068280500006481

Mayabi TW (2019) An artificial neural network model for predicting retail maize prices in Kenya. Ph.D. thesis, University of Nairobi

McIntosh CS, Bessler DA (1988) Forecasting agricultural prices using a Bayesian composite approach. J Agric Appl Econ 20:73–80. https://doi.org/10.1017/S0081305200017611

Melo Bd, Milioni AZ, Nascimento Júnior CL (2007) Daily and monthly sugar price forecasting using the mixture of local expert models. Pesquisa Operacional 27:235–246. https://doi.org/10.1590/S0101-74382007000200003

Minot N (2014) Food price volatility in Sub-Saharan Africa: has it really increased? Food Policy 45:45–56. https://doi.org/10.1016/j.foodpol.2013.12.008

Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw 6:525–533. https://doi.org/10.1016/S0893-6080(05)80056-5

Moreno RS, Salazar OZ et al (2018) An artificial neural network model to analyze maize price behavior in Mexico. Appl Math 9:473. https://doi.org/10.4236/am.2018.95034

Naveena K, Subedar S et al (2017) Hybrid time series modelling for forecasting the price of washed coffee (Arabica Plantation Coffee) in India. Int J Agr Sci (**ISSN. 0975-3710**)

Paluszek M, Thomas S (2020) Practical MATLAB deep learning: a project-based approach. Apress. https://link.springer.com/content/pdf/10.1007/978-1-4842-5124-9.pdf

Piot-Lepetit I, M'Barek R (2011) Methods to analyse agricultural commodity price volatility. 1–11. https://doi.org/10.1007/978-1-4419-7634-5_1

Qazani MRC, Asadi H, Lim CP, Mohamed S, Nahavandi S (2021) Prediction of motion simulator signals using time-series neural networks. IEEE Trans Aerosp Electron Syst 57:3383–3392. https://doi.org/10.1109/TAES.2021.3082662

Rasheed A, Younis MS, Ahmad F, Qadir J, Kashif M (2021) District wise price forecasting of wheat in pakistan using deep learning. arXiv preprint arXiv:2103.04781

Ribeiro MHDM, dos Santos Coelho L (2020) Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. Appl Soft Comput 86:105837. https://doi.org/10.1016/j.asoc.2019.105837

Ribeiro CO, Oliveira SM (2011) A hybrid commodity price-forecasting model applied to the sugar-alcohol sector. Aust J Agric Resour Econ 55:180–198. https://doi.org/10.1111/j.1467-8489.2011.00534.x

Ribeiro MHDM, Ribeiro VHA, Reynoso-Meza G, dos Santos Coelho L (2019) Multi-objective ensemble model for short-term price forecasting in corn price time series. In: 2019 international joint conference on neural networks (IJCNN). IEEE, pp 1–8. https://doi.org/10.1109/IJCNN.2019.8851880

Rl M, Mishra AK (2021) Forecasting spot prices of agricultural commodities in India: application of deep-learning models. Intell Syst Account Financ Manag 28:72–83. https://doi.org/10.1002/isaf.1487

Selvamuthu D, Kumar V, Mishra A (2019) Indian stock market prediction using artificial neural networks on tick data. Financ Innov 5:16. https://doi.org/10.1186/s40854-019-0131-7

Shahhosseini M, Hu G, Archontoulis S (2020) Forecasting corn yield with machine learning ensembles. Front Plant Sci 11:1120. https://doi.org/10.3389/fpls.2020.01120

Shahhosseini M, Hu G, Huber I, Archontoulis SV (2021) Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci Rep 11:1–15. https://doi.org/10.1038/s41598-020-80820-1

Shahwan T, Odening M (2007) Forecasting agricultural commodity prices using hybrid neural networks. In: Computational intelligence in economics and finance. Springer, pp 63–74. https://doi.org/10.1007/978-3-540-72821-4_3

Shimizu S, Kano Y (2008) Use of non-normality in structural equation modeling: application to direction of causation. J Stat Plan Inference 138:3483–3491. https://doi.org/10.1016/j.jspi.2006.01.017

Shimizu S, Hoyer PO, Hyvärinen A, Kerminen A, Jordan M (2006) A linear non-Gaussian acyclic model for causal discovery. J Mach Learn Res 7:2003–2030. https://www.jmlr.org/papers/volume7/shimizu06a/shimizu06a.pdf?ref=https://codemonkey.link

Shimizu S, Inazumi T, Sogawa Y, Hyvärinen A, Kawahara Y, Washio T, Hoyer PO, Bollen K (2011) DirectLiNGAM: a direct method for learning a linear non-gaussian structural equation model. J Mach Learn Res 12:1225–1248. https://www.jmlr.org/papers/volume12/shimizu11a/shimizu11a.pdf

Silalahi DD et al (2013) Application of neural network model with genetic algorithm to predict the international price of crude palm oil (CPO) and soybean oil (SBO). In: 12th national convention on statistics (NCS), Mandaluyong City, Philippine, October, pp 1–2

Silva N, Siqueira I, Okida S, Stevan SL, Siqueira H (2019) Neural networks for predicting prices of sugarcane derivatives. Sugar Tech 21:514–523. https://doi.org/10.1007/s12355-018-0648-5

Storm H, Baylis K, Heckelei T (2020) Machine learning in agricultural and applied economics. Eur Rev Agric Econ 47:849–892. https://doi.org/10.1093/erae/jbz033

Sulewski T, Sprigs J, Schoney R (1994) Agricultural producer price expectations. Can J Agric Econ/Revue canadienne d'agroeconomie 42:301–310. https://doi.org/10.1111/j.1744-7976.1994.tb00026.x

Surjandari I, Naffisah MS, Prawiradinata MI (2015) Text mining of twitter data for public sentiment analysis of staple foods price changes. J Ind Intell Inf. https://doi.org/10.12720/jiii.3.3.253-257

Vajda S, Santosh KC (2016) A fast k-nearest neighbor classifier using unsupervised clustering. In: International conference on recent trends in image processing and pattern recognition. Springer, pp 185–193. https://doi.org/10.1007/978-981-10-4859-3_17

Wang Z, Bessler DA (2004) Forecasting performance of multivariate time series models with full and reduced rank: an empirical examination. Int J Forecast 20:683–695. https://doi.org/10.1016/j.ijforecast.2004.01.002

Wang T, Yang J (2010) Nonlinearity and intraday efficiency tests on energy futures markets. Energy Econ 32:496–503. https://doi.org/10.1016/j.eneco.2009.08.001

Wegener C, von Spreckelsen C, Basse T, von Mettenheim H-J (2016) Forecasting government bond yields with neural networks considering cointegration. J Forecast 35:86–92. https://doi.org/10.1002/for.2385

Wells J, Slade P (2021) The effect of the Canada–China canola trade dispute on canola prices. Can J Agric Econ/Revue canadienne d'agroeconomie 69:141–149. https://doi.org/10.1111/cjag.12258

Wen G, Ma B-L, Vanasse A, Caldwell CD, Earl HJ, Smith DL (2021) Machine learning-based canola yield prediction for site-specific nitrogen recommendations. Nutr Cycl Agroecosyst 121:241–256. https://doi.org/10.1007/s10705-021-10170-5

Xu X (2014a) Causality and price discovery in US Corn Markets: an application of error correction modeling and directed acyclic graphs. Technical Report. https://doi.org/10.22004/ag.econ.169806

Xu X (2014b) Cointegration and price discovery in US Corn Markets. In: Agricultural and resource economics seminar series. North Carolina State University. https://doi.org/10.13140/RG.2.2.30153.49768

Xu X (2014c) Price discovery in US corn cash and futures markets: the role of cash market selection. Technical Report. https://doi.org/10.22004/ag.econ.169809

Xu X (2015a) Causality, price discovery, and price forecasts: evidence from US corn cash and futures markets

Xu X (2015b) Cointegration among regional corn cash prices. Econ Bull 35:2581–2594. http://www.accessecon.com/Pubs/EB/2015/Volume35/EB-15-V35-I4-P259.pdf

Xu X (2017a) Short-run price forecast performance of individual and composite models for 496 corn cash markets. J Appl Stat 44:2593–2620. https://doi.org/10.1080/02664763.2016.1259399

Xu X (2017b) The rolling causal structure between the Chinese Stock Index and futures. Fin Mark Portfolio Mgmt 31:491–509. https://doi.org/10.1007/s11408-017-0299-7

Xu X (2017c) Contemporaneous causal orderings of US corn cash prices through directed acyclic graphs. Empir Econ 52:731–758. https://doi.org/10.1007/s00181-016-1094-4

Xu X (2018a) Causal structure among US corn futures and regional cash prices in the time and frequency domain. J Appl Stat 45:2455–2480. https://doi.org/10.1080/02664763.2017.1423044

Xu X (2018b) Cointegration and price discovery in US corn cash and futures markets. Empir Econ 55:1889–1923. https://doi.org/10.1007/s00181-017-1322-6

Xu X (2018c) Linear and nonlinear causality between corn cash and futures prices. J Agric Food Ind Organ 16:20160006. https://doi.org/10.1515/jafio-2016-0006

Xu X (2018d) Intraday price information flows between the CSI300 and futures market: an application of wavelet analysis. Empir Econ 54:1267–1295. https://doi.org/10.1007/s00181-017-1245-2

Xu X (2018e) Using local information to improve short-run corn price forecasts. J Agric Food Ind Organ. https://doi.org/10.1515/jafio-2017-0018

Xu X (2019a) Contemporaneous and granger causality among US corn cash and futures prices. Eur Rev Agric Econ 46:663–695. https://doi.org/10.1093/erae/jby036

Xu X (2019b) Price dynamics in corn cash and futures markets: cointegration, causality, and forecasting through a rolling window approach. Fin Markets Portfolio Mgmt 33:155–181. https://doi.org/10.1007/s11408-019-00330-7

Xu X (2019c) Contemporaneous causal orderings of CSI300 and futures prices through directed acyclic graphs. Econ Bull 39:2052–2077. http://www.accessecon.com/Pubs/EB/2019/Volume39/EB-19-V39-I3-P192.pdf

Xu X (2020) Corn cash price forecasting. Am J Agric Econ 102:1297–1320. https://doi.org/10.1002/ajae.12041

Xu X, Thurman W (2015a) Forecasting local grain prices: an evaluation of composite models in 500 corn cash markets. https://doi.org/10.22004/ag.econ.205332

Xu X, Thurman WN (2015b) Using local information to improve short-run corn cash price forecasts. https://doi.org/10.22004/ag.econ.285845

Xu X, Zhang Y (2021a) Network analysis of corn cash price comovements. Mach Learn Appl 6:100140. https://doi.org/10.1016/j.mlwa.2021.100140

Xu X, Zhang Y (2021b) Individual time series and composite forecasting of the Chinese Stock Index. Mach Learn Appl 5:100035. https://doi.org/10.1016/j.mlwa.2021.100035

Xu X, Zhang Y (2021c) House price forecasting with neural networks. Intell Syst Appl 12:200052. https://doi.org/10.1016/j.iswa.2021.200052

Xu X, Zhang Y (2021d) Second-hand house price index forecasting with neural networks. J Prop Res. https://doi.org/10.1080/09599916.2021.1996446
Xu X, Zhang Y (2021e) Rent index forecasting through neural networks. J Econ Stud. https://doi.org/10.1108/JES-06-2021-0316
Xu X, Zhang Y (2021f) Corn cash price forecasting with neural networks. Comput Electron Agric 184:106120. https://doi.org/10.1016/j.compag.2021.106120
Xu X, Zhang Y (2022a) Network analysis of housing price comovements of a hundred Chinese cities. Natl Inst Econ Rev. https://doi.org/10.1017/nie.2021.34
Xu X, Zhang Y (2022b) Contemporaneous causality among one hundred Chinese cities. Empir Econ. https://doi.org/10.1007/s00181-021-02190-5
Xu X, Zhang Y (2022c) Soybean and soybean oil price forecasting through the nonlinear autoregressive neural network (NARNN) and NARNN with exogenous inputs (NARNN-X). Intell Syst Appl 13:200061. https://doi.org/10.1016/j.iswa.2022.200061
Xu X, Zhang Y (2022d) Residential housing price index forecasting via neural networks. Neural Comput Appl. https://doi.org/10.1007/s00521-022-07309-y
Xu X, Zhang Y (2022e) Coking coal futures price index forecasting with the neural network. Miner Econ. https://doi.org/10.1007/s13563-022-00311-9
Xu X, Zhang Y (2022f) Thermal coal price forecasting via the neural network. Intell Syst Appl. https://doi.org/10.1016/j.iswa.2022.200084
Xu X, Zhang Y (2022g) Contemporaneous causality among residential housing prices of ten major Chinese cities. Int J Hous Mark Anal. https://doi.org/10.1108/IJHMA-03-2022-0039
Xu X, Zhang Y (2022h) Commodity price forecasting via neural networks for coffee, corn, cotton, oats, soybeans, soybean oil, sugar, and wheat. Intell Syst Account Financ Manag. https://doi.org/10.1002/isaf.1519
Xu X, Zhang Y (2022i) Cointegration between housing prices: evidence from one hundred Chinese cities. J Prop Res. https://doi.org/10.1080/09599916.2022.2114926
Xu X, Zhang Y (2022j) Forecasting the total market value of a shares traded in the Shenzhen Stock Exchange via the neural network. Econ Bull
Xu X, Zhang Y (2022k) Network analysis of price comovements among corn futures and cash prices. J Agric Food Ind Organ. https://doi.org/10.1515/jafio-2022-0009
Xu X, Zhang Y (2022l) Retail property price index forecasting through neural networks. J Real Estate Portfolio Manag. https://doi.org/10.1080/10835547.2022.2110668
Yaakob Z, Mohammad M, Alherbawi M, Alam Z, Sopian K (2013) Overview of the production of biodiesel from waste cooking oil. Renew Sustain Energy Rev 18:184–193. https://doi.org/10.1016/j.rser.2012.10.016
Yang J, Awokuse TO (2003) Asset storability and hedging effectiveness in commodity futures markets. Appl Econ Lett 10:487–491. https://doi.org/10.1080/1350485032000095366
Yang J, Leatham DJ (1998) Market efficiency of US grain markets: application of cointegration tests. Agribus Int J 14:107–112. https://doi.org/10.1002/(SICI)1520-6297(199803/04)14:2107::AID-AGR33.0.CO;2-6
Yang J, Haigh MS, Leatham DJ (2001) Agricultural liberalization policy and commodity price volatility: a Garch application. Appl Econ Lett 8:593–598. https://doi.org/10.1080/13504850010018734
Yang J, Zhang J, Leatham DJ (2003) Price and volatility transmission in international wheat futures markets. Ann Econ Financ 4:37–50. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.295.2182&rep=rep1&type=pdf
Yang J, Su X, Kolari JW (2008) Do euro exchange rates follow a martingale? Some out-of-sample evidence. J Bank Financ 32:729–740. https://doi.org/10.1016/j.jbankfin.2007.05.009
Yang J, Cabrera J, Wang T (2010) Nonlinearity, data-snooping, and stock index ETF return predictability. Eur J Oper Res 200:498–507. https://doi.org/10.1016/j.ejor.2009.01.009
Yang J, Li Z, Wang T (2021) Price discovery in Chinese agricultural futures markets: a comprehensive look. J Future Mark 41:536–555. https://doi.org/10.1002/fut.22179
Yitzhaki S, Slemrod J (1991) Welfare dominance: an application to commodity taxation. Am Econ Rev 480–496
Yoosefzadeh-Najafabadi M, Earl HJ, Tulpan D, Sulik J, Eskandari M (2021) Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean. Front Plant Sci 11:2169. https://doi.org/10.3389/fpls.2020.624273
Yu T-HE, Bessler DA, Fuller SW (2006) Cointegration and causality analysis of world vegetable oil and crude oil prices. Technical Report. https://doi.org/10.22004/ag.econ.21439
Yuan CZ, San WW, Leong TW (2020) Determining optimal lag time selection function with novel machine learning strategies for better agricultural commodity prices forecasting in Malaysia. In: Proceedings of the 2020 2nd international conference on information technology and computer communications, pp 37–42. https://doi.org/10.1145/3417473.3417480
Zelingher R, Makowski D, Brunelle T (2021) Assessing the sensitivity of global maize price to regional productions using statistical and machine learning methods. Front Sustain Food Syst 5:171. https://doi.org/10.3389/fsufs.2021.655206
Zelingher R, Makowski D, Brunelle T (2020) Forecasting impacts of agricultural production on global maize price
Zhang H, Wang Q, Mortimer SR (2012) Waste cooking oil as an energy resource: review of Chinese policies. Renew Sustain Energy Rev 16:5225–5231. https://doi.org/10.1016/j.rser.2012.05.008
Zhang H, Li L, Zhou P, Hou J, Qiu Y (2014) Subsidy modes, waste cooking oil and biofuel: policy effectiveness and sustainable supply chains in China. Energy Policy 65:270–274. https://doi.org/10.1016/j.enpol.2013.10.009
Zhang J, Meng Y, Wei J, Chen J, Qin J (2021) A novel hybrid deep learning model for sugar price forecasting based on time series decomposition. Math Probl Eng. https://doi.org/10.1155/2021/6507688
Zhao H (2021) Futures price prediction of agricultural products based on machine learning. Neural Comput Appl 33:837–850. https://doi.org/10.1007/s00521-020-05250-6
Zou H, Xia G, Yang F, Wang H (2007) An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. Neurocomputing 70:2913–2923. https://doi.org/10.1016/j.neucom.2007.01.009