JOURNAL OF APPLIED ECONOMICS
2019, VOL. 22, NO. 1, 468–483
https://doi.org/10.1080/15140326.2019.1668664

ARTICLE

OPEN ACCESS

Agricultural commodity futures prices prediction via long- and short-term time series network

Hongbing Ouyang[a], Xiaolu Wei[a] and Qiufeng Wu[b]

[a]School of Economics, Huazhong University of Science and Technology, Wuhan, Hubei, P.R. China; [b]School of Science, Northeast Agricultural University, Harbin, Heilongjiang, P.R. China

ABSTRACT
In this study, we attempt to predict global agricultural commodity futures prices through analysis of multivariate time series. Our motivation is based on the notion that datasets of agricultural commodity futures prices involves a mixture of long- and short-term information, linear and non-linear structure, for which traditional approaches such as Auto-Regressive Integrated Moving Average (ARIMA) and Vector Auto-Regression (VAR) may fail. To tackle this issue, Long- and Short-Term Time-series Network (LSTNet) is applied for prediction. Empirical results show that LSTNet achieves better performance over that of several state-of-the-art baseline methods on average and in most periods based on three performance evaluation measures and two tests of performance difference.

ARTICLE HISTORY
Received 13 May 2019
Accepted 12 September 2019

KEYWORDS
Agricultural commodity futures; prices prediction; long- and short-term timeseries network; long-term information

1. Introduction

Forecasting agricultural commodity futures prices is an important subject in agricultural domain, not only in providing price information of agricultural commodity in advance which decision makers rely on, but also reducing the uncertainty and risks of agricultural markets (Wang, Yue, Wei, & Lv, 2017). Futures prices are also used by crop insurance programs to decide their first-stage and harvest prices (Zulauf, Rettig, Roberts, & Matt, 2015). However, agricultural commodity futures characteristics, noisy and non-stationary, make prediction face challenges (Xiong, Li, Bao, Hu, & Zhang, 2015). “Noisy” implies that there is insufficient information to observe past behaviors of agricultural commodity futures. “Non-stationary” means that agricultural commodity futures may change dramatically in different periods. These characteristics lead to poor agricultural commodity futures’ prediction results as predicted by traditional econometric models such as linear model, Auto-Regressive Integrated Moving Average (ARIMA) and Vector Auto-Regression (VAR) (Onour & Sergi, 2011; Zulauf, Irwin, Ropp, & Sberna, 1999.; Zafeiriou & Sariannidis, 2011). The aforementioned methods are generally based on the assumption that variables are independent, normal distribution, which is contradicted with real market.

In recent years, time series prediction based on neural networks has been popular. Unlike traditional models, deep neural networks have several distinct advantages as

CONTACT Xiaolu Wei xiaoluees@hust.edu.cn School of Economics, Huazhong University of Science and Technology, Wuhan, Hubei, P.R. China

non-parametric, self-learning, non-assumption and noise tolerant, which are unavailable in traditional models (Haykin & Network, 2004; Hochreiter & Schmidhuber, 1997; Sharda & Patil, 1992). Therefore, deep neural networks might be more effective in forecasting agricultural commodity futures price in comparison to traditional models (Kaastra & Boyd, 1995; Zhang & Hu, 1998).

However, machine learning techniques, which are seldom used in commodity futures prices prediction, especially agricultural commodity futures prices. In this study, we attempt to predict global agricultural commodity futures prices through Long- and Short-term Time Series Network (LSTNet) developed by Guokun Lai et al. (Lai, Chang, Yang, & Liu, 2018). Based on the notion that governments usually control the price and quantity of agricultural commodities, price of agricultural commodities is more susceptible to government policies, which have a time lag. Therefore, datasets of agricultural commodity futures prices might involve a mixture of extreme long- and short-term information, LSTNet method that could capture such information might be more accurate in the prediction of agricultural commodity futures prices.

The rest of this article is organized as follows. Section 2 summarizes the literature review of time series prediction in machine learning. Section 3 describes the mathematical model on agricultural commodity futures price prediction and detailed introduction to LSTNet method. Section 4 describes the empirical preliminaries, which contain empirical dataset and selection of evaluation criteria. Section 5 presents the empirical steps of LSTNet method and the empirical results. Finally, the conclusion is drawn in Section 6.

2. Literature review

Since the first study of Tomek and Gray (1970), the predictive performance of agricultural futures has been the focus of scholars' in-depth study. Based on empirical forecast assessment, they found that futures prices are good price predictors in the corn and soybean market. However, the results of many subsequent studies on different agricultural futures markets have been mixed. This is partly due to the highly dependency of forecasting performance on specific market conditions and traditional econometric models, which are usually used by researchers (Garcia & Leuthold, 2004; Kenyon, Jones, & McGuirk, 1993; Kofi, 1973; Zulauf et al., 1999). Moreover, there mainly exist two problems in literature related to agricultural futures prediction through traditional econometric models. One is that high dimensional multivariate time series prediction is rarely analyzed due to the model capacity and their high computational cost (Zafeiriou & Sariannidis, 2011; Zulauf et al., 1999). The second one is that these models are generally based on the assumption that variables are independent, normal distribution, which is unrealistic in the real market (Onour & Sergi, 2011).

Recently, deep neural networks provide a promising tool in time series forecasting (Adya & Collopy, 1998; Tang, De Almeida, & Fishwick, 1991) due to its ability to model nonlinear pattern, realize complex causal relationship, and learn from huge history dataset. In the field of time series prediction through deep neural networks, there exist various approaches, such as long and short-term memory (LSTM) (Jia, 2016) and support vector machine (SVM) (Tay & Cao, 2001). The studies related to time series prediction through deep neural network mainly have three categories. One is to identify statistically significant events in time series (Chau & Wong, 1999; Liu & Yue, 2018;

Malhotra, Vig, Shroff, & Agarwal, 2015; Namaki, Lin, & Wu, 2017). The second one is to seek and predict inherent structure in the time series. The third one is to predict numerical value of time series accurately.

However, the performance of machine learning in agricultural commodity futures prediction is rarely explored (Xiong et al., 2015). The few studies related to agricultural commodity futures prediction through machine learning mainly focus on interval forecasting of agricultural commodity futures prices. Moreover, there exist four problems in agricultural commodity futures prediction through neural networks. One is that existing studies mainly focus on interval prediction while ignore the point forecasting of agricultural commodity futures prediction in neural networks fields. The second one is that existing few studies mainly ignore the dynamic dependencies among multiple variables. The third one is that existing methods fail to capture very long-term information due to gradient vanishing. The forth one, which is also the most important one, is that these studies fall short in distinguishing a mixture of short-term and long-term repeating patterns explicitly (Cao & Tay, 2003; Connor, Atlas, & Martin, 1992; Dasgupta & Osogami, 2017).

The contributions of this article are, firstly, to fill the gap between agricultural commodity futures prices' point forecasting and machine learning techniques, and secondly, to predict agricultural commodity futures prices simultaneously with full consideration of interaction among variables, which is helpful to make investment portfolio of agriculture commodity futures, and thirdly, to capture extremely long-term and short-term information of agricultural commodity futures prices with Recurrent-skip module in the LSTNet method, and fourthly, to measure relative forecast performance of LSTNet method and other methods on average as well as in dynamic environment. Therefore, we can predict agricultural commodity futures prices more accurately through the application of LSTNet method and verify relative prediction performance more efficiently with several robust tests.

3. Agricultural commodities futures price prediction based on Lstnet

3.1. Mathematical model on agricultural commodities futures price prediction

This article focuses on forecasting multivariate agricultural commodity futures prices with long-term and short-term repeating patterns. Dataset on agricultural commodity futures prices mainly has two characteristics: correlation and a combination of short and long-term repeating patterns. Since agricultural commodities change collaboratively due to weather, market, and other conjunctural factors, dataset on agricultural commodity futures prices consists of 12 variables, including CZCE cotton, CZCE sugar, ICE eleventh sugar, DCE bean, DCE bean II, DCE soybean oil, DCE cardamom, CZCE strong wheat, DCE corn, ICE coffee, ICE cocoa, ICE frozen concentrated orange juice. Moreover, the subsequences of these variables represent short-term and long-term repeating patterns. This problem could be solved by constructing convolutional neural network (CNN) module and Recurrent-skip component, which are introduced by LSTNet method. CNN module is able to extract short-term patterns in the time dimension as well as local dependencies between variables, while Recurrent-skip component could capture extremely long length of one pattern with temporal skip-connections. Therefore, we could discover short-term and long-term repeating patterns of multivariate agricultural commodity futures prices and predict prices more accurately.

This article aims at predicting agricultural multivariate commodity futures prices. Given datasets $X = \{x_t\}_{t=1}^{T}$, wherein $x_t \in R^n$ and n is the variable dimension, this article is interested in the task of forecasting a series of agricultural commodity futures prices in a rolling forecasting step. Instead of looking at a single independent variable y_t, this article predicts x_{T+h} with n dimensions simultaneously, wherein h is the desirable horizon ahead of the current time stamp and $\{x_t\}_{t=1}^{T}$ are available. Similarly, to forecast x_{T+h+1} in the next time step and assume $\{x_t\}_{t=1}^{T+1}$ are available. Therefore, the input matrix of this article at time stamp T is $X = \{x_t\}_{t=1}^{T} \in R^{n\times T}$, and the output matrix is $x_{T+h} \in R^n$.

3.2. LSTNet on agricultural commodities futures price prediction

Neural network is widely used for time series prediction. In the prediction process, three problems must be considered. The first one is the consideration of very long-term information and the combination of short-term and very long-term repeating patterns. Many studies prefer time series prediction of short-term repeating patterns rather than a combination of short-term and very long-term repeating patterns, which is used by the method applied in this article. The second one is the ignorance of linear and non-linear structures, which usually yield unsatisfactory outcomes. The last problem is the efficiency of the forecasting process. With the increase of data size, the time needed to predict time series increases remarkably. Therefore, an algorithm to reduce the total number of data points and the operation time of time series prediction is indispensable. These three problems are closely related to time series prediction process.

LSTNet was proposed by Guokun Lar et al. in 2018, which was accepted by the International Conference on Special Interest Group on Information Retrieval (SIGIR). Compared to other forecasting methods, LSTNet is the first method to predict n dimensional time series with a mixture of short-term and extremely long-term repeating patterns, which could solve the above problems. In this section, we describe the details of LSTNet algorithm applied in this article.

LSTNet consists of a non-linear part and a linear part. The non-linear part includes a convolutional layer, a dense layer consists of recurrent component and recurrent-skip component, a temporal attention layer, while the linear part uses a autoregressive model (AR) to forecast the result.

3.2.1. Convolutional layer

The first layer of LSTNet is a convolutional network. Given the input matrix X, this section extracts short-term patterns and interdependences among 12 variables. The output in this section can be expressed as,

$$h_k = \mathrm{RELU}(W_k * X + b_k) \tag{1}$$

where w denotes the number of width filters, n is the number of variables which is set to be 12 in this article, * is the convolution operation, and the RELU function is RELU(x) = max(0,x).

3.2.2. Recurrent component and recurrent-skip component

Given the output of convolutional layer, these recurrent component and recurrent-skip component aims at capturing long-term and very long-term information, the outputs of recurrent layer and recurrent-skip layer are the hidden states at each time stamp. The hidden states of recurrent layer's units at time t can be formulated as,

$$
\begin{aligned}
r_t &= \sigma(x_t W_{xr} + h_{t-1} W_{hr} + b_r) \\
u_t &= \sigma(x_t W_{xu} + h_{t-1} W_{hu} + b_u) \\
c_t &= RELU(x_t W_{xc} + r_t \odot (h_{t-1} W_{hc}) + b_c) \\
h_t &= (1 - u_t) \odot h_{t-1} + u_t \odot c_t)
\end{aligned}
\tag{2}
$$

where $\odot$ is the element-wise product, σ is the sigmoid function, and x_t is the input of recurrent layer as well as the output of convolutional layer at time stamp t.

However, due to gradient vanishing, the Recurrent layers with GRU (Chung, Gulcehre, Cho, & Bengio, 2014) in this article may fail to capture very long-term correlation in real world. LSTNet method proposes a recurrent structure with temporal skip-connections to investigate history information in a longer time period, namely Recurrent-Skip Component. The updating process can be expressed as,

$$
\begin{aligned}
r_t &= \sigma\left(x_t W_{xr} + h_{t-p} W_{hr} + b_r\right) \\
u_t &= \sigma\left(x_t W_{xu} + h_{t-p} W_{hu} + b_u\right) \\
c_t &= RELU\left(x_t W_{xc} + r_t \odot (h_{t-p} W_{hc}\right) + b_c) \\
h_t &= \left(1 - u_t\right) \odot h_{t-p} + u_t \odot c_t)
\end{aligned}
\tag{3}
$$

where p is the number of hidden cells skipped through which is determined empirically in this article.

Given outputs of the Recurrent component at time t and Recurrent-skip component from time t-p + 1 to t, denoted by $\left\{h_t^R, h_{t-p+1}^S, h_{t-p+2}^S \cdots, h_t^S\right\}$, this article uses a dense layer to produce the prediction result of the LSTNet's non-linear part which can be expressed as,

$$
h_t^D = W^R h_t^R + \sum\nolimits_{i=0}^{p-1} W_i^S h_{t-i}^S + b \tag{4}
$$

3.2.3. Temporal attention layer

However, the Recurrent-skip layer component maybe unfavorable in the non-seasonal time series prediction or repeating patterns forecasting with flexible time period. To alleviate this problem, LSTNet develops an attention mechanism (Bahdanau, Cho, & Bengio, 2014). The output of this temporal attention layer is a non-linear projection part, which is computed as,

$$h_t^D = W\left[c_t; h_{t-1}^R\right] + b \tag{5}$$

where h_{t-1}^R is last window hidden state, $c_t = H_t\alpha_t$ is the weighted context of hidden states of the input matrix, α_t is the attention weights which can be expressed as,

$$\alpha_t = \mathrm{AttnScore}\left(H_t^R, h_{t-1}^R\right) \tag{6}$$

3.2.4. Autoregressive layer

Due to non-linear property of Convolutional layer and Recurrent layer, LSTNet method decomposes the prediction into a non-linear part, which is captured by Convolutional layer and Recurrent layer, and a linear part, which is solved by Autoregressive (AR) model in this section. Given the initial input X, we can get the forecasting result of the linear part through AR Layer, which is formulated as follows,

$$h_{t,i}^L = \sum\nolimits_{k=0}^{q^{ar-1}} W_k^{ar} y_{k-1,i} + b^{ar} \tag{7}$$

Then, the forecasting result of LSTNet can be expressed as follows,

$$\hat{Y}_t = h_t^D + h_t^L \tag{8}$$

The pseudo code of Toeplitz Graphical Lasso is described in Table 1.

Table 1. LSTNet framework.

Algorithm 1 LSTNet framework
Input initial $X = \{x_t\}_{t=1}^T$, wherein $x_t \in R^n$ $(n = 12)$
Output a mixed output $\hat{Y}_t$ of a linear part h_t^L and a non-linear part h_t^D
Initialize best_val = 10,000,000
for i ← 1 to epoches do
$h_k \leftarrow \mathrm{RELU}(W_k * X + b_k)$
if p > 0 do
$h_t \leftarrow (1 - u_t) \odot h_{t-1} + u_t \odot c_t)$
$h_t \leftarrow (1 - u_t) \odot h_{t-p} + u_t \odot c_t)$
$h_t^D \leftarrow W^R h_t^R + \sum_{i=0}^{p-1} W_i^S h_{t-i}^S + b$
else
$h_t^D \leftarrow W\left[c_t; h_{t-1}^R\right] + b$
if highway > 0 do
$h_{t,i}^L \leftarrow \sum_{k=0}^{q^{ar-1}} W_k^{ar} y_{k-1,i} + b^{ar}$
$\hat{Y}_t \leftarrow h_t^D + h_t^L$
if val_loss < best_val
best_val = val_loss
model save
else
continue
end

4. Data

4.1. Data set and data preprocessing

The selection of agricultural commodity futures and time period are based on a trade-off between more variables and longer price histories, which might be helpful in the training of LSTNet method. However, the LSTNet method applied in this article has been well verified in the datasets of a relatively short time period and a long time period. Therefore, 12 agricultural commodity futures collated from Wind platform are analyzed in the empirical section. They are CZCE cotton (CF.CZC), CZCE sugar (SR. CZC), ICE sugar (SB.NYB), DCE bean (A.DCE), DCE bean II (B.DCE), DCE soybean oil (Y.DCE), DCE cardamom (M.DCE), CZCE strong wheat (WH.CZC), DCE corn (C. DCE), ICE coffee (KC.NYB), ICE cocoa (CC.NYB), ICE frozen concentrated orange juice (0J.NYB). The data set cover the time period from 01/09/2006 up to 01/10/2019.

More specifically, closing prices are used as datasets, which are illustrated in Figure 1. In order to have a better understanding of historical variation, (a), (b), (c) show agricultural commodity futures prices in daily, monthly, and yearly scale, respectively. Moreover, Figure (d) shows the daily closing prices of agricultural commodity futures

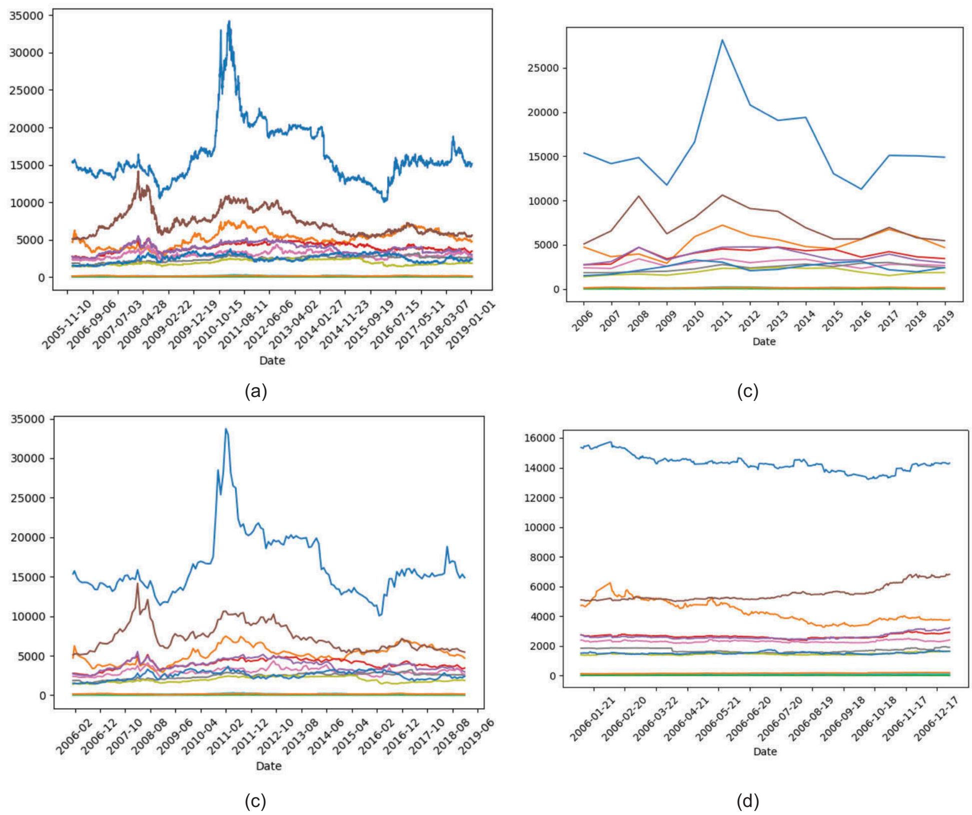

Figure 1. Closing prices of agricultural commodity futures prices.
Source: Wind, 2006–2019

prices during 2006. The short-term and long-term repeating patterns are not clear due to nonstational time series or patterns with flexible time period. Each sample data of agricultural commodity futures prices is split into training set (60%), validation set (20%), test set (20%) in chronological order. The study uses validation set to tune hyper parameters, while uses test set to evaluate and compare forecasting performance of LSTNet and other models. In addition, the Null values are immediately dropped due to its little scale.

4.2. Unit root analysis

Agricultural commodity futures have characteristics of noisy and non-stationary, which may make prediction results unsatisfied (Xiong et al., 2015). Therefore, this section implements a comprehensive unit root test to analyze the stationarity of agricultural commodity futures, which include Levin-Lin-Chu test (LLC), Im-Pesaran-Shin test (IPS) and Phillips-Perron test (PP). The results of statistical value are found in Table 2.

Table 2 shows the stationary test results of agricultural commodity futures. Clearly, with the null hypothesis that assumes nonstationary, agricultural futures are not stationary while the first-order difference of all agricultural commodity futures is stationary. Therefore, in order to compare the forecasting results of LSTNet method and other techniques, such as CNN and RNN, and verify the performance of LSTNet method, we will apply these techniques in datasets with no difference and first order difference.

4.3. Performance criteria

The prediction performance of LSTNet method is compared with CNN, RNN, ARIMA, VAR. CNN, RNN, VAR are able to analyze multivariate input and output, while ARIMA is a single output method in which we will train n models independently. Here, n is the number of variables in datasets, which is set to be 12 in this article. In order to verify the validity of LSTNet method proposed in this article, we select five evaluation methods, including three performance measures and two tests of performance differences. Performance measures include Root Relative Squared Error (RSE), Relative Absolute Error (RAE), Empirical Correlation Coefficient (CORR). The RSE and RAE are in scaled version, which are designed to make comparisons more efficient and valid. Tests of performance differences include a multistep conditional predictive ability test proposed by Giacomini and White (2006) and a fluctuation test proposed by

Table 2. *Unit root tests for* agricultural commodity futures.

Method	Statistic	Prob.**
No difference		
LLC	−0.07687	0.4694
IPS	−0.26162	0.3968
PP	22.2885	0.5620
First order difference		
LLC	−364.143	0.0000
IPS	−239.637	0.0000
PP	2812.80	0.0000

Giacomini and Rossi (2010). Multistep conditional predictive ability test is a performance test on average, while fluctuation test is a performance test to evaluate the time-varying relative forecast performance of the models. For tests of relative forecast performance, we just sum the RSE and RAE of the 12 output variables at each time point and reject hypothesis that two models have equal out-of-sample performance when test statistic is larger than the critical value. The definitions of these criteria are found in Table 3.

Here, $\mathrm{Y}, \widehat{\mathrm{Y}} \in \mathrm{R}^{\mathrm{n}\times\mathrm{T}}$ are true values and predicted values of agricultural futures prices, respectively, n is the number of out of sample forecasts, τ is forecast horizon, T is total sample size, m is maximum estimation of window size, h_t is a test function, ΔL_i is out-of-sample forecast loss differences of two methods, $\bar{Z}_{m,n} = n^{-1}\sum_{t=m}^{T-\tau} Z_{m,t+\tau}$, $Z_{m,t+\tau} = h_t \Delta L_{m,t+\tau}$, $\tilde{\Omega}_n = n^{-1} Z_{m,t+\tau} Z'_{m,t+\tau} + n^{-1}\sum_{j=1}^{\tau-1} w_{n,j} \times \sum_{t=m+j}^{T-\tau}\left[Z_{m,t+\tau}Z'_{m,t+\tau-j} + Z_{m,t+\tau-j}Z'_{m,t+\tau}\right]$, where $w_{n,j}$ is a weight function (West, 1987).

5. Results and discussion

To predict agricultural commodity futures prices, we apply LSTNet method to multivariate time series prediction of agricultural commodity futures with no difference and first order difference. LSTNet uses five components to extract short-term and long-term repetitive patterns with consideration of linear and non-linear structure of time series. These five components are convolutional component, recurrent component, recurrent-skip component, temporal attention component, and autoregressive component. We repeat the algorithm until we find the lowest validation loss value. The results and discussion of the empirical research is described in the following section.

In the prices prediction of agricultural commodity futures based on LSTNet, a skip length p of 24 is found to produce the best possible results. The hidden dimension of Recurrent and Convolutional layer, the dropout rate, the horizon h and the optimization algorithm are arbitrarily chosen to be 50, 0.2, 12, and the Adam algorithm, respectively. The program is constructed using python 3 language. Table 4 shows the performance comparison of agricultural commodity futures prices prediction on average, while relative performances based on MSE and MAE at each point with a horizon

Table 3. Performance criteria and their calculations.

Criteria	Calculation
RSE	$\mathrm{RSE} = \frac{\sqrt{\sum_{(i,t)\in\Omega_{Test}}\left(Y_{it}-\hat{Y}_{it}\right)^2}}{\sqrt{\sum_{(i,t)\in\Omega_{Test}}\left(Y_{it}-\mathrm{mean}(Y)\right)^2}}$
RAE	$\mathrm{RAE} = \frac{\sum_{(i,t)\in\Omega_{Test}}\left\lvert Y_{it}-\hat{Y}_{it}\right\rvert}{\sum_{(i,t)\in\Omega_{Test}}\left\lvert Y_{it}-\mathrm{mean}(Y)\right\rvert}$
CORR	$\mathrm{CORR} = \frac{1}{n}\sum_{i=1}^{n}\frac{\sum_t\left(Y_{it}-\mathrm{mean}(Y_i)\right)\left(\hat{Y}_{it}-\mathrm{mean}\left(\hat{Y}_i\right)\right)}{\sqrt{\sum_t\left(Y_{it}-\mathrm{mean}(Y_i)\right)^2\left(\hat{Y}_{it}-\mathrm{mean}\left(\hat{Y}_i\right)\right)^2}}$
Multistep Conditional Predictive Ability Test	$T^h_{m,n,\tau} = n\left(n^{-1}\sum_{t=m}^{T-\tau} h_t\Delta L_{m,t+\tau}\right)\tilde{\Omega}_n^{-1}\left(n^{-1}\sum_{t=m}^{T-\tau} h_t\Delta L_{m,t+\tau}\right) = n\bar{Z}'_{m,n}\tilde{\Omega}_n^{-1}\bar{Z}_{m,n}$
Fluctuation test	$F^{OOS}_{t,m} = \hat{\sigma}^{-1}m^{-1/2}\sum_{j=t-m/2}^{t+\frac{m}{2}-1}\Delta L_j$

Table 4. Performance comparison of different methods.

Methods	Metrics	Horizon 3	6	9	12	15	18	21	24
LSTNet-Skip	RSE	0.0373	0.0533	0.0691	0.0741	**0.0342**	0.0905	**0.0414**	0.0908
	RAE	0.0279	0.0423	0.0585	0.0616	**0.0328**	0.0727	**0.04**	0.0771
	CORR	0.9385	0.8975	0.8736	0.8225	0.7654	0.7442	0.7155	0.6871
LSTNet-Attn	RSE	**0.0347***	**0.0444***	**0.0551***	**0.0636***	0.0678*	**0.0721***	0.0821*	**0.0831***
	RAE	**0.0258***	**0.0325***	**0.0396***	**0.0461***	0.0506*	**0.0561***	0.0624*	**0.0632***
	CORR	**0.9475**	**0.9135**	0.8748	0.835	0.8074	**0.7808**	**0.746**	0.7238
RNN	RSE	0.0395	0.0522	0.0578	0.0703	0.0704	0.0784	0.0855	0.0889
	RAE	0.0323	0.0425	0.0452	0.0566	0.0559	0.0613	0.0667	0.0703
	CORR	0.9474	0.9119	**0.8761**	**0.8351**	**0.8136**	0.777	0.7448	**0.7257**
CNN	RSE	0.0415	0.0514	0.0701	0.0738	0.1077	0.0895	0.1037	0.1061
	RAE	0.0297	0.04	0.053	0.0565	0.0898	0.0694	0.0896	0.0954
	CORR	0.9323	0.8922	0.8001	0.8162	0.7582	0.717	0.6898	0.6558
ARIMA-stationary	RSE	1.0018	1.0006	1.0021	1.0029	1.0021	1.0020	1.0022	1.0013
	RAE	1.0015	1.0004	1.0010	1.0013	0.9998	0.9997	1.0010	0.9990
	CORR	−0.0474	−0.0411	−0.0446	−0.0442	−0.0417	−0.0394	−0.0395	−0.0327
VAR-stationary	RSE	1.0002	1.0002	1.0003	1.0002	1.0002	1.0002	1.0002	1.0002
	RAE	1.0000	0.9996	0.9996	0.9994	0.9982	0.9981	0.9991	0.9975
	CORR	−0.0378	−0.0513	−0.0572	−0.0531	−0.0520	−0.0483	−0.0483	−0.0442
LSTNet-Skip-stationary	RSE	1.0065	1.0143	1.0175	1.0207	1.01	1.0224	1.0231	1.0092
	RAE	1.0252	1.0278	1.0255	1.0346	1.026	1.0343	1.0331	2.0196
	CORR	0.0194	0.0075	0.0009	0.0136	−0.0005	−0.0157	−0.0071	0.0219
LSTNet-Attn-stationary	RSE	1.0049	1.0027	1.0069	1.0065	1.0084	1.0092	1.0068	1.0084
	RAE	1.0096	1.0067	1.0092	1.0075	1.0109	1.011	1.0079	1.0103
	CORR	0.001	−0.0021	−0.0035	−0.0052	−0.0023	−0.0107	−0.0031	−0.0144
RNN-stationary	RSE	1.0088	1.0046	1.0074	1.0074	1.0076	1.0097	1.0134	1.0083
	RAE	1.0172	1.0094	1.0095	1.0095	1.0123	1.0146	1.0161	1.0122
	CORR	−0.0009	−0.0083	−0.0028	−0.0028	0.0186	0.0001	−0.0338	−0.012
CNN-stationary	RSE	1.0204	1.0022	1.0105	1.0202	1.0144	1.0093	1.021	1.0175
	RAE	1.0302	1.0155	1.0225	1.0291	1.0286	1.0175	1.028	1.033
	CORR	−0.0052	−0.01	0.0078	−0.0176	0.0004	−0.0073	−0.0158	0.0034

of 6 are reported in Figures 2 and 3, respectively. Moreover, main diagnostic tests of ARIMA and VAR are reported in Appendix A and B, the numbers within parentheses are p-value.

Table 4 summarizes the prediction performances of all methods on all test sets (20%) in all metrics, including RSE, RAE, CORR of LSTNet-Skip, LSTNet-Attn, RNN, CNN, ARIMA-stationary, VAR-stationary, LSTNet-Skip-stationary, LSTNet-Attn-stationary, RNN-stationary, CNN-stationary. The data for the last six models is

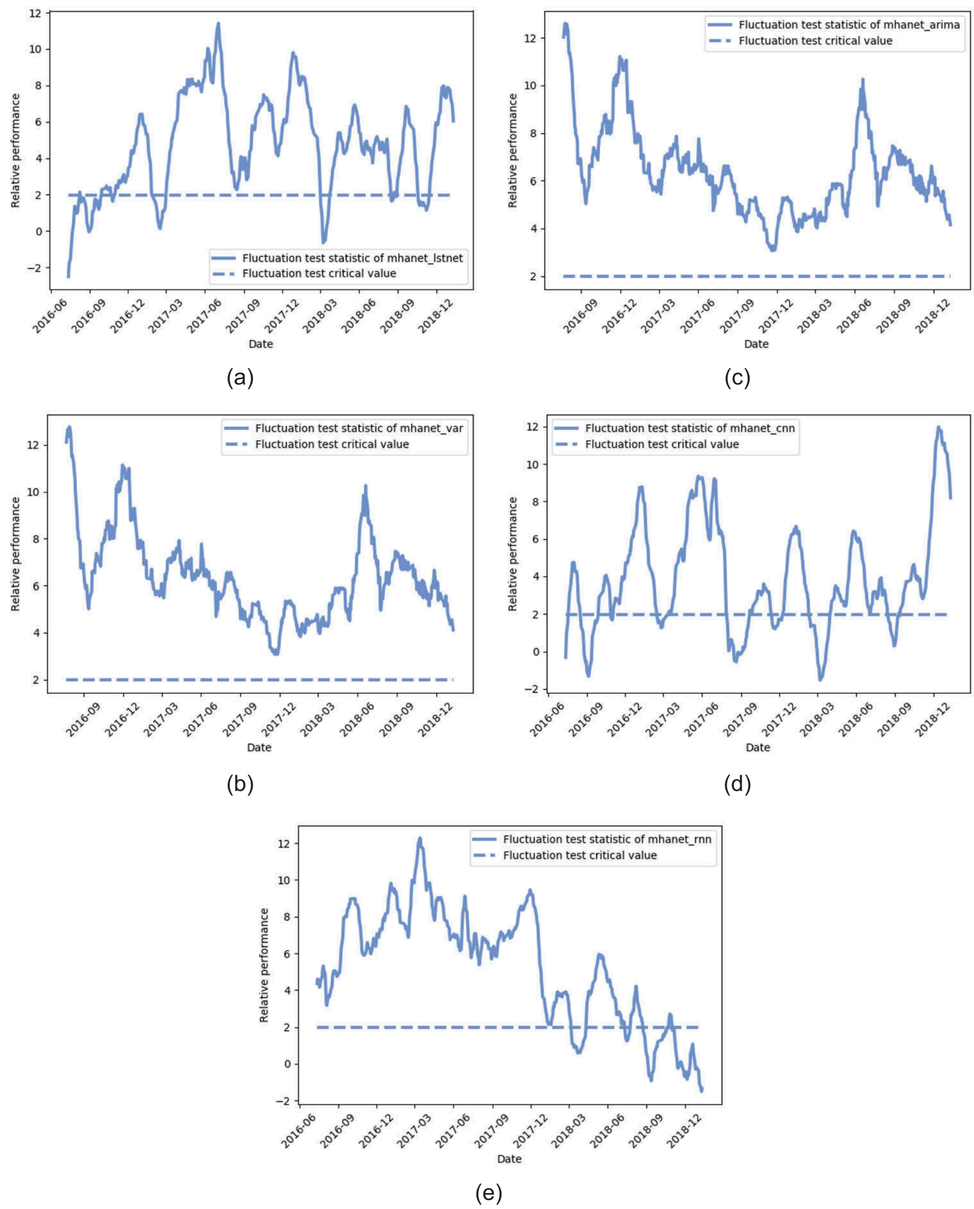

Figure 2. Relative performance based on MSE in dynamic environment.

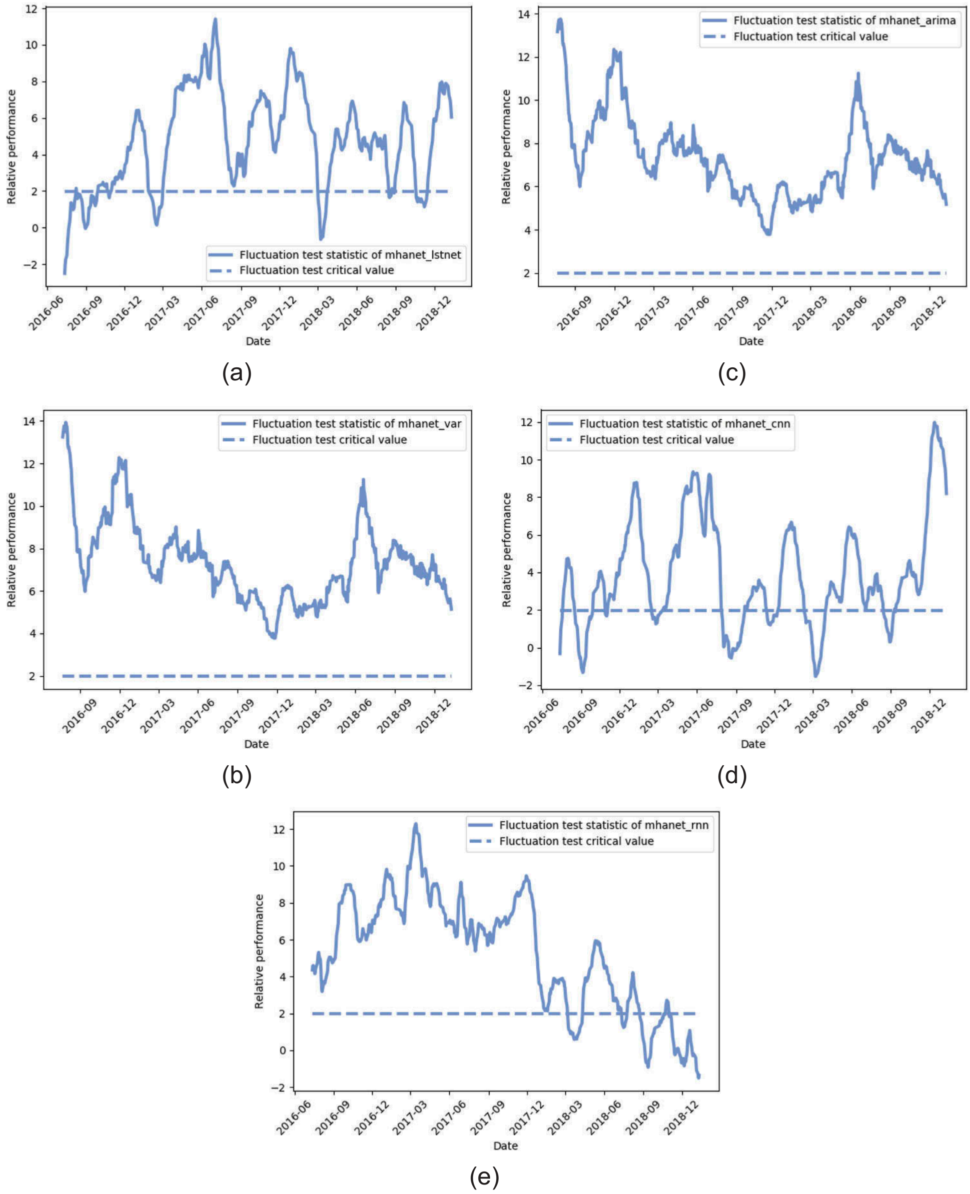

Figure 3. Relative performance based on MAE in dynamic environment.

agricultural commodity futures prices after first-order difference. Moreover, we set horizon = {3, 6, 9, 12, 15, 18, 21, 24}, respectively. The larger the horizons, the worse the prediction results. The best result for four method and three metrics is highlighted in bold face in Table 4. The total count of the bold-faced results is 16 for LSTNet-Attn, 4 for LSTNet-Skip, 4 for RNN, and 0 for CNN. Moreover, an asterisk sign (*) indicates that the test rejects equal conditional predictive ability at the 1% level and that the LSTNet-Attn method outperforms other methods through conditional predictive ability tests on average.

Clearly, even though the periodic patterns of global agricultural commodity futures prices are not clear and the dataset is nonstationary, LSTNet still perform better than other neural network methods (RNN, CNN) and traditional econometric methods (ARIMA, VAR) on average. Specifically, LSTNet-Attn outperforms the neural baseline RNN, CNN, ARIMA, VAR by 6.52, 21.68, 91.70,91.69% in RSE metric and 10.10, 33.75, 93.67, 93.66% in RAE metric respectively when the horizon is 24, suggesting much better performance of the proposed method.

Figures 2 and 3 report the results for the GW Fluctuation test in the horizon of 6 which measures relative forecast performances between LSTNet-Attn and other methods in dynamic environment. The figures report both the Fluctuation test statistic as well as the one-sided critical value at 5%. It is obvious that LSTNet-Attn has better forecast performance over ARIMA and VAR. Moreover, the figures point out that there have been more periods in which LSTNet-Attn performs better than other methods.

The LSTNet-Attn method has robust performance in different metrics, partly due to its consideration of interdependencies among multiple variables, extremely long-term information and linear structure.

6. Conclusions

Time series prediction with neural networks provides a fundamental aid for the comprehension of global agricultural commodity futures prices and, more specifically, is crucial in reducing uncertainty and risks in agricultural markets. In the literature of multivariate time series prediction through neural networks, they mainly focus on forecasting univariate time series without considering interdependencies among different variables, and moreover, generally fail to capture very long-term information and linear structure. In this article we apply a rather different approach: Long- and Short-Term Time-series network (LSTNet) to predict prices of several global agricultural commodity futures. LSTNet method is a new prediction model that is able to predict multivariate time series simultaneously with full consideration of a mixture of extremely long-term and short-term patterns, linear and non-linear structures.

This study applies the LSTNet method to forecast agricultural commodity futures prices simultaneously. LSTNet method consists of five components, including CNN component, RNN component, RNN-skip component, Temporal Attention component, and Autoregressive component. The empirical research shows that by combing the strengths of convolutional network, recurrent network, and autoregressive component, LSTNet method significantly improves the state-of-the-art results in multivariate time series forecasting on the dataset of agricultural commodity futures prices. With the empirical results, we show the applied LSTNet method is a promising alternative for multivariate time series forecasting in agricultural commodity futures markets.

There are two promising extensions of multivariate time series prediction in agricultural commodity futures prices. A possible extension of agricultural commodity futures prices prediction is to investigate possibility of performance improvements in multivariate time series forecasting when the datasets represent unobvious repetitive patterns. The other is to analyze automatic adjustment of hyperparameters, including recurrent-skipped numbers p and horizon h, which are tuned manually in LSTNet.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Project ID: 19YJA790067).

Notes on contributors

Hongbing Ouyang was born at Hubei Province, China, in 1968. He received PhD in Finance from City University of Hong Kong, in November, 2004. He is now a professor in School of Economics, Huazhong University of Science and Technology. His research interest includes financial economics, financial econometrics, investments and financial regulation.

Xiaolu Wei was born at Hubei Province, China, in 1991. She received Master degree in mathematical economics from Northeastern Agricultural University, Harbin, China, in June 2016. She is studying for a PhD degree in finance from Huazhong University of Science and Technology, Wuhan, China. She has passed Level III examination of CFA in 2016. Her current research interests include financial application, economic prediction and corporate finance.

Qiufeng Wu was born at Heilongjiang Province, China, in 1979. He received Ph.D in computer application technology from Harbin Institute of Technology, Harbin, China, in January 2014. He is working as associated professor in College of Science in Northeast Agricultural University. He is CCF member. His current research interests include machine learning, computer vision and smart agriculture.

References

Adya, M., & Collopy, F. (1998). How effective are neural networks at forecasting and prediction? A review and evaluation. *Journal of Forecasting*, *17*(5-6), 481–495.

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv Preprint arXiv:1409.0473*.

Cao, L. J., & Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, *14*(6), 1506–1518.

Chau, T., & Wong, A. K. (1999). Pattern discovery by residual analysis and recursive partitioning. *IEEE Transactions on Knowledge and Data Engineering*, *11*(6), 833–852.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv Preprint arXiv:1412.3555*.

Connor, J., Atlas, L. E., & Martin, D. R. (1992). Recurrent networks and NARMA modeling. In *Advances in neural information processing systems*.(pp. 301–308).

Dasgupta, S., & Osogami, T. (2017, February). Nonlinear dynamic Boltzmann machines for time-series prediction. In *Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, California*.

Garcia, P., & Leuthold, R. M. (2004). A selected review of agricultural commodity futures and 17 options markets. *European Review of Agricultural Economics*, *31*(3), 235–272.

Giacomini, R., & Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, *25*(4), 595–620.

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*, 1545–1578.

Haykin, S., & Network, N. (2004). A comprehensive foundation. *Neural Networks*, *2*(2004), 41.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Jia, H. (2016). Investigation into the effectiveness of long short term memory networks for stock price prediction. *arXiv Preprint arXiv:1603.07893*.

Kaastra, I., & Boyd, M. S. (1995). Forecasting futures trading volume using neural networks. *Journal of Futures Markets*, *15*(8), 953–970.

Kenyon, D., Jones, E., & McGuirk, M. A. (1993). Forecasting performance of corn and soybean harvest futures contracts. *American Journal of Agricultural Economics*, *75*(2), 399–407.

Kofi, T. A. (1973). A framework for comparing the efficiency of futures markets. *American Journal of Agricultural Economics*, *55*(4), 584–594.

Lai, G., Chang, W. C., Yang, Y., & Liu, H. (2018, June). Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*(pp. 95–104). Ann Arbor, MI: ACM.

Liu, D., & Yue, S. (2018). Event-driven continuous STDP learning with deep structure for visual pattern recognition. *IEEE Transactions on Cybernetics*, (99), 1–14.

Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015, April). Long short term memory networks for anomaly detection in time series. *The 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Belgium*, 89.

Namaki, M. H., Lin, P., & Wu, Y. (2017, December). Event pattern discovery by keywords in graph streams. In *2017 IEEE International Conference on Big Data (Big Data)* (pp. 982–987), Boston, MA, USA. IEEE.

Onour, I., & Sergi, B. S. (2011). Modeling and forecasting volatility in the global food commodity prices (Modelování a Prognózování Volatility Globálních cen Potravinářských Komodit). *Agricultural Economics-Czech*, *57*(3), 132–139.

Orzechowski, P., Sipper, M., Huang, X., & Moore, J. H. (2018). EBIC: An artificial intelligence-based parallel biclustering algorithm for pattern discovery. *arXiv Preprint, arXiv:1801.03039*.

Sharda, R., & Patil, R. B. (1992). Connectionist approach to time series prediction: An empirical test. *Journal of Intelligent Manufacturing*, *3*(5), 317–323.

Tang, Z., De Almeida, C., & Fishwick, P. A. (1991). Time series forecasting using neural networks vs. *Box-Jenkins Methodology. Simulation*, *57*(5), 303–310.

Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *omega*, *29*(4), 309–317.

Tomek, W. G., & Gray, R. W. (1970). Temporal relationships among prices on commodity futures markets: Their allocative and stabilizing roles. *American Journal of Agricultural Economics*, *52*(3), 372–380.

Wang, D., Yue, C., Wei, S., & Lv, J. (2017). Performance analysis of four decomposition-ensemble models for one-day-ahead agricultural commodity futures price forecasting. *Algorithms*, *10*(3), 108.

West, N. K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–708.

Xiong, T., Li, C., Bao, Y., Hu, Z., & Zhang, L. (2015). A combination method for interval forecasting of agricultural commodity futures prices. *Knowledge-Based Systems*, *77*, 92–102.

Zafeiriou, E., & Sariannidis, N. (2011). Nonlinearities in the price behaviour of agricultural products: The case of cotton. *International Journal of Food, Agriculture & Environment-JFAE*, *9*(2), 551–555.

Zhang, G., & Hu, M. Y. (1998). Neural network forecasting of the British pound/US dollar exchange rate. *Omega*, *26*(4), 495–506.

Zulauf, C. R., Irwin, S. H., Ropp, J. E., & Sberna, A. J. (1999). A reappraisal of the forecasting performance of corn and soybean new crop futures. *Journal of Futures Markets*, *19*(5), 603–618.

Zulauf, Rettig, C., Roberts, N., & Matt. (2015). Do futures forecast the future? *Farmdoc Daily*, (4), 147.

Appendix A Diagnostic test of ARIMA*

	horizon							
Residual Analysis	3	6	9	12	15	18	21	24
Durbin Watson	2.0910	2.0984	2.0916	2.0910	2.1284	2.1119	2.0813	2.1346
Normality_Jarque Bera	372.0859	384.7351	388.2332	400.9036	421.2390	441.8660	462.9237	486.6546
	(1.59E-81)	(2.86E-84)	(4.97E-85)	(8.81E-88)	(3.38E-92)	(1.12E-96)	(3E-101)	(2.11E-106)
Normality_Omni	67.8933	68.8597	69.0394	69.8428	73.0176	76.5639	77.7150	84.6265
	(1.81E-15)	(1.12E-15)	(1.02E-15)	(6.82E-16)	(1.39E-16)	(2.37E-17)	(1.33E-17)	(4.20E-19)

[a]The numbers within parentheses are p-value.

APPENDIX B Diagnostic test of var*

	horizon							
Residual Analysis	3	6	9	12	15	18	21	24
Durbin Watson	2.0971	2.1013	2.0986	2.0940	2.1302	2.1142	2.0830	2.1394
Normality_Jarque Bera	382.0133	388.8561	388.0492	402.9957	421.8263	442.8802	467.3759	491.2635
	(1.11E-83)	(3.64E-85)	(5.45E-85)	(3.09E-88)	(2.52E-92)	(6.76E-97)	(3.24E-102)	(2.11E-107)
Normality_Omni	67.78028	68.76885	68.46677	69.85638	72.86031	76.55144	77.641017	84.97134
	(1.91E-15)	(1.17E-15)	(1.36E-15	(6.77E-16)	(1.51E-16)	(2.38E-17)	(1.38E-17)	(3.54E-19)

[a]The numbers within parentheses are p-value.