Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 162 (2019) 33–38

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Information Technology and Quantitative Management (ITQM 2019)

Carbon futures price forecasting based with ARIMA-CNN-LSTM model

Lei Ji [a,b], Yingchao Zou [c], Kaijian He [a,b,*], Bangzhu Zhu [d*]

[a] *School of Business, Hunan University of Science and Technology, Xiangtan,411201,China*
[b] *Hunan Engineering Research Center for Industrial Big Data and Intelligent Decision Making, Hunan University of Science and Technology, Xiangtan 411201, China*
[C] *College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China*
[d] *Business School, Nanjing University of Information Science & Technology, Nanjing 210044, China*

Abstract

In this paper, we introduced an ARIMA-CNN-LSTM model to forecast the carbon futures price. The ARIMA-CNN-LSTM model employs the ARIMA model and the deep neural network structure that combines the CNN and LSTM layers to capture linear and nonlinear data features. In ARIMA-CNN-LSTM model structure, the ARIMA is used to capture the linear features. The Convolutional Neural Network (CNN) is used to capture the hierarchical data structure while the Long Short Term Memory network (LSTM) is used to capture the long-term dependencies in the data. Comprehensive performance evaluation has been conducted using weekly carbon futures price. Results have confirmed that ARIMA-CNN-LSTM model can achieve better prediction accuracy than the benchmark models, in terms of Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) performance measures.

Keywords: Carbon futures price forecasting; deep learning model; ARIMA model; Convolutional Neural Network (CNN); Long Short-Term Memory Network (LSTM)

1. Introduction

The rapid development of the world economy has been accompanied by the increasing level of greenhouse gas emissions worldwide [1]. Greenhouse gases emitted by industrial production is closely linked to global climate change and sustainable development. As the climate issues such as global warming have demonstrated significant negative impact on the development of human society and global environmental quality, researchers as well as the government and industries pay more attention to the techniques to deal with these arising problems

* Corresponding author. Tel.: +86-731-58290306; fax:+86-731-58290306.
E-mail address: kaijian.he@my.cityu.edu.hk.

10.1016/j.procs.2019.11.254

such as energy conservation, emission reduction and the greenhouse gas emission [2][3]. Since the enforcement of the Kyoto protocol in 1997, carbon trading and management is viewed now as one of the key tools to mitigate the effects of global climate change [4]. The launch of European Union Emissions Trading Scheme (EU ETS) in 2005 has accelerated the development of the global carbon trading market ever since [3].

As the carbon market trading becomes more and more active, the accurate forecasting of carbon futures price has attracted a lot of attention in the research community. In the literature, main approaches have covered the traditional linear time series models to the more recent artificial intelligence approaches. For example, [5] proposed a hybrid forecasting model that combined the traditional ARIMA model and the least squares support vector machine (LSSVM). Experimental results showed that the proposed hybrid forecasting model has produced better prediction performance in term of the carbon futures price forecasting. [6] proposed a hybrid prediction model that combined the multi-output support vector regression (MSVR) and particle swarm optimization (PSO) to forecast the price of carbon future. The researchers found that the proposed hybrid forecasting model have achieved the best forecasting performance on the carbon futures price. [7] proposed a hybrid forecasting model that combined the variational mode decomposition (VMD) and spiking neural networks (SNNs). The researchers found that the proposed VMD-SNN forecasting model produced a better prediction performance on the prediction of carbon futures price. In the meantime, new models such as the CNN, LSTM, LightGBM emerging in the engineering field have demonstrated great potential in forecasting the market price movement in the finance and economics field [8]. The convolutional neural network (CNN) has the capability to capture the important spatial futures in the carbon futures price by convolution operation and pooling operation [9]. The long short-term memory network (LSTM) can learn the long-time dependencies in the time series and retain information span long time [10]. Therefore, the deep neural network models combining CNN and LSTM have been widely used in different disciplines. For example,.[11] used the recurrent neural networks (RNN) and long short-term memory network (LSTM) and convolutional neural network (CNN) to forecast the stock price. They found that the prediction ability of convolutional neural network (CNN) has achieved better forecasting performance than the ARIMA model. [12] combined the deep recurrent neural network (RNN) with deep convolutional neural network (CNN) to form a hybrid forecasting method, and then used the proposed hybrid to forecast the foreign exchange time series data. In their research, they used different financial varieties as the input data to forecast the closing price. [13] used different structures of LSTM model to forecast the weather. Through the experimental results, they found that the multi-layer LSTM model have better forecasting accuracy.[14] proposed a forecasting model that combined the hybrid CNN-LSTM model and the dense layer to predict the electricity load. They used the hybrid CNN-LSTM model to capture the features of the historical load and used the dense layer to capture the features of other correlated variables, and then forecast the load according to these extracted features.[15] proposed a prediction model combining the 2D CNN model and LSTM model to make prediction on traffic. [16] proposed a hybrid forecasting model that combined the one-dimensional CNN model and LSTM model to predict the concentration of PM2.5. They used three factors, such as cumulated hours of rain, cumulated wind speed and historical PM2.5 concentration, to make prediction on the PM2.5 concentration. However, majority of CNN-LSTM models has been designed to extract multivariate data features from the multivariate data modeling. Very few have been applied to time series forecasting, not to mention the analysis and modeling of carbon futures prices.

The contribution of this paper is that we introduced the hybrid forecasting model with ARIMA model and the deep neural network model combining the one-dimensional CNN and LSTM layers to forecast carbon futures price. We used the ARIMA model to model the linear features of carbon future price and used the one-dimensional CNN model to extract the spatial characteristic of the residual of ARIMA model and then used the LSTM model to capture the long-term dependencies of features over the time. More specifically, the one-dimensional CNN model is employed to reduce the influence of white noise by extracting important features in time series and LSTM model is employed to capture the long-term dependencies of these extracted features. In this paper, we have evaluated the forecasting performance of a wide range of models such as LSTM model, CNN model and the hybrid ARIMA-CNN-LSTM model, against the forecasting performance of the benchmark ARIMA model. The experiment results show that the ARIMA-CNN-LSTM forecasting model has achieved the best prediction accuracy in term of carbon futures price forecasting.

The rest of the paper consists of the following sections. Section 2 presents the detailed ARIMA-CNN-LSTM model. Section 3 reports and analyzes the empirical results using the carbon futures price. Section 4 provides the concluding summary.

2. Methodology

ARIMA-CNN-LSTM model combined linear ARIMA model, the CNN model and the LSTM model. Firstly, we used the ARIMA model to forecast the carbon future price and calculated the residual of the ARIMA model. Then we assume that there exist spatial relationship temporal dependencies between adjacent observations of the residual of ARIMA model. We take advantage of the unique spatial and temporal modelling capability of CNN model and LSTM layers to extract spatio-temporal features in the residual of ARIMA model. More specifically, we used CNN model as the filter to extract the spatial relationship between neighbouring residual, and then used LSTM model to capture the long-term temporal dependencies of these features extracted by CNN model [14][15].The general structure of the hybrid CNN-LSTM model is illustrated in figure 1.

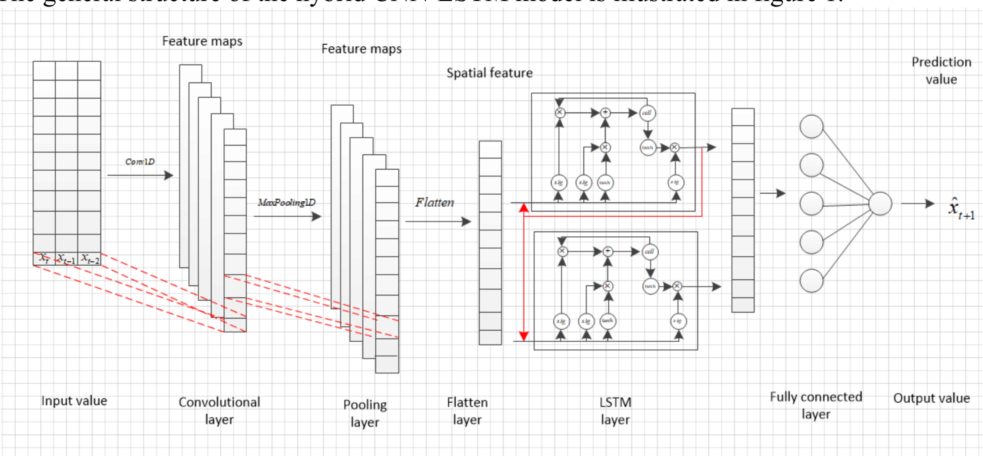

Fig. 1. the structure of the hybrid CNN-LSTM model

As shown in figure 1, CNN-LSTM model consists of input layer, convolutional layer, pooling layer, flatten layer, LSTM layer and fully connected layer. The convolutional layer extracts the spatial features from the carbon futures price at adjacent time point by convolution operation and output the feature maps. The convolutional neural network is the popular deep learning model that is frequently used in the area of image recognition [11]. The advantage of the CNN is that it can capture the spatial features in the observations. The one-dimensional convolutional neural network is used to extract the distinctive space relationship features existing in the neighboring residual of ARIMA model. The typical CNN model has five different sections that are input layer, convolutional layer, pooling layer, fully connected layer and output layer [12]. The units of the convolutional layer locally connect to the input layer. Each unit in the convolutional layer has the same weights matrix connection [17], which can decrease the parameters of CNN and make the training of CNN model become easy. The convolutional layer captures the features in the input signal through the convolution operations. The pooling layer is used to down sample the output feature maps and it can summarize the most important features of extracted feature maps by convolutional layer [18]. The flatten layer transforms these feature maps into one-dimensional vectors as the input into the LSTM layer.

The LSTM layer models the long-term temporal dependencies in these feature maps. The fully connected layer makes prediction according to these extracted features. The long short-term memory network (LSTM) is a specific structure of the recurrent neural network (RNN), and is typically employed to capture the long-term dependencies of the time series [19]. The special architecture of the long short-term memory network (LSTM) can solve the gradients vanished problem existing in the traditional recurrent neural network (RNN) when modeling the long-term dependencies in the time series[20][21]. The most important structure of the LSTM is the memory block, which is contained in hidden layer. There are one or more memory cells, an input gate, a forget gate and a output gate be contained in the memory block. The memory cell is a self-connection unit and it is capable storing the temporal state of the LSTM network [22][19]. The three gates are the especial multiplicative units, which are respectively used to control the input, save and output of the flowing information in the memory block of LSTM.

3. Empirical Studies

In this research, we used the weekly futures price of the carbon in the EU emission trading system as the main research dataset. The research dataset can be obtained from the website https://sandbag.org.uk/carbon-price-viewer/. The dataset covers the period from 7 April 2008 to 6 May 2019. The dataset contains a total of 573 weekly observations. In this study, we divided the dataset into training set and testing set. The training set contains 400 observations covering about the first 70% of the raw dataset and the testing set includes 173 observations covering about the last 30% of the raw dataset. In this study, the software that we used are python and matlab, the packages in the software including Keras, numpy, tensorflow, pandas, etc. [23]. The specification of the hardware device we used is that Intel Core i7-7700HQ, NVIDIA GeForce GTX1060, 64-bit Windows10 Operating System and RAM 16GB, etc. The main prediction accuracy evaluation criteria are the mean absolute percentage error (MAPE) and the root mean square error (RMSE) [22].

Firstly, we used the Augmented Dickey-Fuller (ADF) Test to test the stationarity of the raw dataset. The null hypothesis of the ADF test is not rejected. Therefore, the raw dataset needs to transform into a stationary dataset through the first order difference operation before we used the ARIMA model. Then we found that the ARIMA (1,1,1) model have the minimum value of AIC among different model specifications. We used the ARIMA (1,1,1) model as benchmark model to forecast carbon futures price. The number of predicted values is 173. The LSTM model has 1 hidden layer and the hidden layer have 54 hidden neurons. In the training process, the rolling window was set as 3 and the rolling window moves forward one step each time. And the training Epochs is 250 and the initial learning rate is 0.01. In the test process, we used the trained LSTM model to make one-step ahead forecasting with the testing set . In the CNN forecasting process, we set the rolling window of CNN model equal to 3. The CNN model has 1 convolutional layer, 1 pooling layer and 1 fully connected layer, the convolutional layer has 128 filters and the activation function of the convolutional layer is Rectified Linear Unit(ReLU), the pooling approach of the pooling layer is max pooling and the pooling size is equal to 2.In the prediction process of hybrid ARIMA-CNN-LSTM model, firstly, we used the ARIMA model to capture the linear features of carbon future price and calculated the residual of ARIMA model, and then used the hybrid CNN-LSTM model to capture the spatio-temporal features of ARIMA residual. We assumed that there exists the close relationship between adjacent ARIMA model residual, therefore, we used the CNN model to capture the spatial relationship of carbon futures price residual for three consecutive weeks. So, the size of the rolling window we set is 3, the CNN-LSTM model has 1 convolutional layer,1 pooling layer,1 LSTM layer and 1 fully connected layer. The convolutional layer has 45 filters, the activation function of the convolutional layer is Rectified Linear Unit(ReLU). The pooling approach of the pooling layer is max pooling and the pooling size is equal to 1. The

LSTM layer has 10 units and the activation function is Rectified Linear Unit (ReLU). And the learning rate of hybrid CNN-LSTM is 0.001, training epochs is 30.

Extensive empirical studies in the energy market show that the energy price is influenced by diverse range of factors such as the investor sentiment, fundamental factor and financial factors [24][25][26]. Some of the factors have impacts over different time horizon and may demonstrate significantly different linear or nonlinear behaviors. As the carbon market is closely related to the energy consumption and energy market, the carbon futures price is also affected by many factors and demonstrates both linear and nonlinear characteristics. We use different forecasting model to capture the different features of carbon futures price and shows the out-of-sample experiment results of different forecasting models in Table 1.

Table 1. mean absolute percentage error (MAPE) and root mean square error (RMSE) of different forecasting models

	CNN	LSTM	ARIMA	ARIMA-CNN-LSTM
RMSE	0.8616	0.7251	0.7015	0.6940
MAPE	0.0623	0.0456	0.0423	0.0421

As shown in Table 1, the MAPE and RMSE of the hybrid ARIMA-CNN-LSTM model are the smallest. When compared with the ARIMA model, the RMSE of hybrid ARIMA-CNN-LSTM model decreased by 1.07% and the value of MAPE decreased by 0.47%. When compared with the CNN model, the RMSE of hybrid ARIMA-CNN-LSTM model decreased by 19.45% and the value of MAPE decreased by 32.42%. When compared with the LSTM model, the RMSE of the hybrid ARIMA-CNN-LSTM model decreased by 4.29% and the value of MAPE decreased by 7.68%. The experimental results indicate that the ARIMA-CNN-LSTM model has achieved the best prediction performance on the carbon futures price forecasting when compared with other forecasting model. Our results indicate that the ARIMA-CNN-LSTM have better ability to capture the spatio-temporal features of carbon futures price.

4. Conclusions

In this paper, we have proposed the ARIMA-CNN-LSTM model to forecast the carbon futures price. We used the ARIMA model to capture the linear data feature of carbon futures price, and used the CNN model to extract the distinctive spatial relationship features of ARIMA residual, and then utilized the LSTM model to capture the long-term temporal dependencies of these features extracted by CNN model. The experimental results demonstrated that the ARIMA-CNN-LSTM model can fit the carbon futures price better. In this carbon futures price forecasting research, the combination of ARIMA model, CNN model and LSTM model can give full play to their respective advantages to improve the prediction ability of individual model applications.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (NSFC no. 71671013), and Humanities and Social Sciences Youth Foundation of the Ministry of Education of China (no. 16YJC790026).

References

[1] Guoxing Zhang, Zhenhua Zhang, Peng Liu, Mingxing Liu. The running mechanism and prediction of the growth rate of China's Carbon emissions, Chinese Journal of Management Science 2015, Volume 23(12), Pages 86-93.

[2] Bangzhu Zhu, Dong Han, Ping Wang, Zhanchi Wu, Tao Zhang, Yi-Ming Wei. Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. Applied Energy 2017,Volume 191, Pages 521-530.

[3] Mohamed El Hédi Arouri, Fredj Jawadi, Duc Khuong Nguyen,.Nonlinearities in carbon spot-futures price relationships during Phase II of the EU ETS. Economic Modelling 2012, Volume 29, Pages 884-892.
[4] Bao-jun Tang, Cheng Shen, Chao Gao, The efficiency analysis of the European CO2 futures market. Applied Energy 2013, Volume 112, Pages 1544-1547.
[5] Bangzhu Zhu, Yiming Wei, Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology, Omega 2013, Volume 41, Pages 517-524.
[6] Lu Zhang, Junbiao Zhang, Tao Xiong, and Chiao Su,Interval Forecasting of Carbon Futures price Using a Novel Hybrid Approach with Exogenous Variables,Discrete Dynamics in Nature and Society, vol. 2017, Article ID 5730295.
[7] Sun, Guoqiang and Chen, Tong and Wei, Zhinong and Sun, Yonghui and Zang, Haixiang and Chen, Sheng, A Carbon Price Forecasting Model Based on Variational Mode Decomposition and Spiking Neural Networks. Energies 2016, Volume 9, Pages 54.
[8] Xiaolei Sun, Mingxi Liu, Zeqian Sima. 2018. A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Research Letters. 2018. Forthcoming, https://doi.org/10.1016/j.frl.2018.12.032
[9] Jesus Lago, Fjo De Ridder, Bart De Schutter, Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms;Applied Energy 2018,Volume 221,386-405.
[10] Sundaravelpandian Singaravel, Johan Suykens, Philipp Geyer,Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction,Advanced Engineering Informatics 2018,Volume 38,81-90.
[11] Hiransha M, Gopalakrishnan E.A., Vijay Krishna Menon, Soman K.P., NSE Stock Market Prediction Using Deep-Learning Models,Procedia Computer Science 2018, Volume 132,Pages 1351-1362.
[12] Lina Ni, Yujie Li, Xiao Wang, Jinquan Zhang, Jiguo Yu, Chengming Qi,Forecasting of Forex Time Series Data Based on Deep Learning,Procedia Computer Science 2019, Volume 147,Pages 647-652.
[13] Afan Galih Salman, Yaya Heryadi, Edi Abdurahman, Wayan Suparta,Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting,Procedia Computer Science 2018, Volume 135,Pages 89-98.
[14] Wan He,Load Forecasting via Deep Neural Networks.Procedia Computer Science,Procedia Computer Science 2017,Volume 122,Pages 308-314.
[15] Haiyang Yu,Zhihai Wu,Shuqin Wang, Yunpeng Wang,Xiaolei Ma,2017,Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks., Sensors 2017,Volume 17 :1501.
[16] Huang, Chiou-Jye and Kuo, Ping-Huan,A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities,Sensors 2018,Volume 18, page 2220.
[17] Kuo, Ping-Huan & Huang, Chiou-Jye.. A High Precision Artificial Neural Networks Model for Short-Term Energy Load Forecasting.Energies 2018,Volume 11, Article ID 2013.
[18] Hu, Baotian and Lu, Zhengdong and Li, Hang and Chen, Qingcai, Convolutional Neural Network Architectures for Matching Natural Language Sentences. Advances in Neural Information Processing Systems 27 (NIPS 2014) 2014, Pages 2042--2050.
[19] Afan Galih Salman, Yaya Heryadi, Edi Abdurahman, Wayan Suparta,Weather Forecasting Using Merged Long Short-Term Memory Model (LSTM) and Autoregressive Integrated Moving Average (ARIMA) Model,Journal of Computer Science 2018, Volume 14, Pages 930-938.
[20] Joerg Evermann, Jana-Rebecca Rehse, Peter Fettke,Predicting process behaviour using deep learning,Decision Support Systems 2017, Volume 100,Pages 129-140.
[21] Thomas Fischer, Christopher Krauss,Deep learning with long short-term memory networks for financial market predictions,European Journal of Operational Research 2018, Volume 270, Pages 654-669.
[22] Alaa Sagheer, Mostafa Kotb,Time series forecasting of petroleum production using deep LSTM recurrent networks,Neurocomputing 2019, Volume 323,Pages 203-213.
[23] Kaijian He, Lei Ji, Geoffrey K.F. Tso, Bangzhu Zhu, Yingchao Zou,Forecasting Exchange Rate Value at Risk using Deep Belief Network Ensemble based Approach.Procedia Computer Science 2018,Volume 139,Pages 25-32.
[24] Qiang Ji, Jianping Li, Xiaolei Sun, Measuring the interdependence between investor sentiment and crude oil returns: New evidence from the CFTC's disaggregated reports. Finance Research Letters. 2019, Volume 30, Pages 420-425.
[25] Yan-ran Ma, Qiang Ji*, Jiaofeng Pan. Oil financialisation and volatility forecast: Evidence from multidimensional predictors. Journal of Forecasting 2019, Volume 38, Pages 564-581.
[26]]Xiaolei Sun, Xiaoyang Yao, Yuying Yang, Dengsheng Wu, Jianping Li. Multiscale feature extraction and identification of country risk dynamics: cases of OPEC countries. Chinese Journal of Management Science 2015, Volume 23, Pages1-10.