Journal of Scientific & Industrial Research
Vol. 79, October 2020, pp. 945-950

A Hybridized Forecasting Model for Metal Commodity Prices: An Empirical Model Evaluation

Nirjharinee Parida[1]*, Debahuti Mishra[1], Kaberi Das[1], Narendra Kumar Rout[2] and Ganapati Panda[3]

[1]Department of Comp. Sc. and Engg., Siksha 'O'Anusandhan Deemed to be University, Odisha 751 030, India

[2]Department of Comp. application, National Institute of Technology, Raipur (C.G) 492 010, India

[3]Department of Electr. & Telecom. Engg., C. V. Raman Global University, Odisha 752 054, India

Received 30 April 2020; revised 21 May 2020; accepted 15 July 2020

Appropriate decision on perfect commodity prediction under market's constant fluctuations intensifies the need for efficient methods. The main objective of this study is to apply different optimization algorithms such as conventional particle swarm optimization (PSO), bat algorithm (BAT) and ant colony optimization (ACO) algorithms on back propagation neural network (BPNN) to enhance the accuracy of prediction and minimize the error. In this paper, a model has been proposed for volatility forecasting using PSO algorithm to train the BPNN and predict the commodities' closing price. The proposed PSO-BPNN model is considered as best forecasting model compared to BPNN, BAT-BPNN and ACO-BPNN. The experiment has been carried out upon five publicly available metal datasets (gold, silver, lead, aluminium, and copper) to forecast the price return volatility of those five metals challenging the effectiveness. Here, three technical indicators and four filters, such as; moving average convergence/divergence (MACD), williams %R (W%), bollinger (B), least mean squares (LMS), finite impulse response (FIR), Kalmanand recursive least square (RLS) have been applied for providing an additional degree of freedom to train and test the classifiers. From the experimental result analysis it has been found that the proposed PSO-BPNN produces promising output while comparing with BPNN, BAT-BPNN and ACO-BPNN.

Keywords: Ant colony optimization, BPNN, Commodity market forecasting, PSO

Introduction

As per the investment concern and focus on future investment, commodity trading is the perfect alternative to expand the investor's portfolios.[1] With the increase of machine learning approaches having the ability of computational price forecasting against frequent volatility[2] on metal commodity sector plays an interesting opportunity to investors.

Quick convergence[3] feature of ant colony optimization (ACO), solving non-linear problems of BAT algorithm[4–6], and convenient performance[7] of particle swarm optimization (PSO) present themselves to work with back propagation neural network (BPNN) against over-fitting dilemmas[8] providing fascinating prediction results. A hybrid model is very much necessary for making perfect prediction on commodity price which obtains the reliable prediction accuracy with superior performance and resolves the limitation of single optimized model.[7,9] Chang and Lee[10] proved that the particle swarm optimization back propagation network (PSOBPN) produces better forecasting optimal solution as compared to genetic algorithm back propagation network (GABPN) predicting NTD/USD exchange rates. The investigation report of Salman *et al.*[11] exhibits accurate forecasting on crude palm oil price with a powerful combination of PSO-BPNN optimization techniques. According to Rehman and Nawi[12], BAT-BPNN algorithm is introduced whose role is to find the optimal weights in BPNN for avoiding its local minima problem.

In forecasting analysis of commodity market, the successful relationship between PSO and BPNN is motivated to propose a commodity forecast model by hybridizing both. Since BPNN has benefited power for solving nonlinear problems[13] with high accuracy. In order to optimize the results in simulation, a series of sensitizations of the BPNN architecture is conducted on *gold*, *silver*, *lead*, *aluminium* and *copper* metals to forecast the closing price and volatility. More experiments were also conducted to show the performance of the proposed hybridized model on different statistics and measures along with

*Author for Correspondence
E-mail: paridanirjharinee@gmail.com

mean squared error (MSE), mean absolute percent error (MAPE), and mean absolute deviation (MAD) to proof the effectiveness of our approach.

Schematic layout of proposed forecasting model with problem formulation section, Let $d_i = \{c_i, o_i, l_i, h_i\}$ be the dataset of n features where, $d_i = \{c_i, o_i, l_i, h_i\}$ representing closing, open, low and high value of i^{th} feature of D and D $\in \mathfrak{R}$. Here, i defines the number of feature of daily indices of commodity market dataset. Objective of prediction model is to determine the value for d_{n+1} when $d_{1..n}$ can be formulated using Eq. (1).

$$c_{i+1} = F((c_i, o_i, l_i, h_i), (c_{i-1}, o_{i-1}, l_{i-1}, h_{i-1}), \dots (c_1, o_1, l_1, h_1)) \quad \dots (1)$$

Where, c_{i+1} is the closing value on $(i+1)^{th}$ day obtained as classification result of i^{th} feature vectors by computing any classifier F. Target variable c_{i+1} is replaced by $(\mu(c_{i+1} - c_i))/(\sigma c_i)$, where, μ and σ are the *mean* and *standard deviation* of dataset D. With this the input to network have been extended by using different filter outputs evaluated on actual and technical (processed) datasets using Eq. (2). Where, τ_i and ϑ_i are the technical indicators and filter response of D.

$$\left(\mu\left(c_{i+1} - c_i\right)\right) \Big/ (\sigma c_i) = F(\tau_i,\ \tau_{i-1}, \dots, \vartheta_i, \vartheta_{i-1}, \dots)$$

Three technical indicators like *moving average convergence/ divergence (MACD)*, *williams %R (W%), bollinger (B),* are used for smoothening the closing price index of different datasets and elevating the forecasting performance of commodity market price trends as well. The four filters i.e., *least mean squares (LMS), finite impulse response (FIR), Kalman and recursive least square (RLS)* are used to reduce the errors with more numbers by filtering the redundant data occurring in the prediction. They have been used to transform four basic daily inputs into eleven correlated attributes thus giving additional degree of freedom to train and test the classifiers.

The architecture of our proposed system is shown in Fig. 1. The basic datasets with four (*close*, *open*, *high* and *low* price) are transformed into eleven features. This feature vector is then used to train and test proposed PSO-BPNN network. In this work, we have used BPNN consisting of input, one hidden, and output layer in the ratio of $[n \times 2n + 1 \times 1]$ respectively, where n is the number of inputs. This network is trained with BPNN algorithm. PSO algorithm is used to optimize the weight of neural network. Classification results are compared with the expected results using Eq. (2), and producing signal one of three values i.e., *Purchase* (P), *Sell* (S), *Maintain* (M). Proposed system produces the signal P when the compared result is greater than threshold; if the compared result is less than negative threshold it produces S, otherwise M. Positive threshold $+\delta$ and negative threshold $-\delta$ are considered for this experimentation is 0.0005 and −0.0005 respectively.

In the training steps of PSO-BPNN, global best (G_{best}) solution is the optimized weight of network. The testing of network which utilizes the G_{best} solution obtained from the training steps as weight of the network. Initially velocity of particle and position of particle are initialized with random number of dimension. Dimension has got the dependency over number of connections in the network among neurons. For experimentation in this work, we have considered three layered architecture with $11\ (inputlayer) \times 15\ (hiddenlayer) \times 1\ (outputlayer)$ number of neurons randomly chosen. In the training

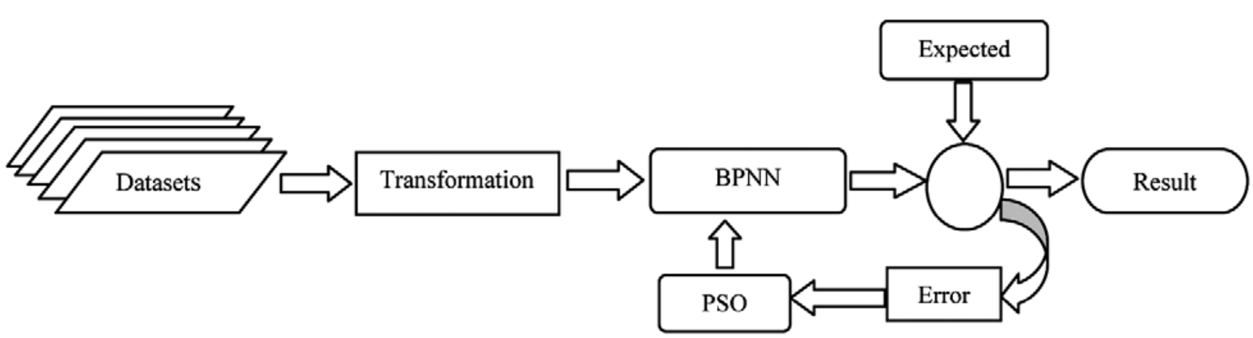

Fig. 1 — Proposed architecture

steps, extraction function is used to extract weight matrix of order $n \times hn$, where n is the number of input is feature and hn is the number of hidden neurons from i^{th} feature of position matrix. Fitness of the network is running sum of error square for each feature vector. The P_{best} and G_{best} is used to update the velocity and position of swarm to be used in next iteration. This process continues till the maximum iteration is achieved.

During testing phase, the $gbest$ position achieved from the training phase has been used, to update the *weight* for network.

From the simulation, the forecasting error of PSO-BPNN is found low indicating better forecast performance compared to ACO-BPNN and BAT-BPNN.

Experimentation and Model Evaluation

This forecasting method has been implemented using MATLAB 10, on a Pentium-4, 4 GB RAM and 2-GHz processor machine under Windows XP platform.

Dataset discussion

This study was carried out using five benchmark datasets and total number of features collected from website *www.forex.com* for all five dataset is 855, 1114, 1113, 983, 1019 respectively over the time period from January 2013 to February 2017. Out of this, the corresponding dimension of 90% training datasets have been taken as 770 × 4, 1003 × 4, 1002 × 4, 885 × 4 and 918 × 4 for Gold, Silver, Lead, Aluminium and Copper. The remaining 10 % of datasets is used for testing purpose. The statistics result where hypothesis test, p-value, confidence interval, t-stat, degree of freedom and standard deviations computed for all five metal datasets are given in Table 1.

Table 1 — T-test statistics of five metal datasets

Dataset	Confidence Interval	Stats		
		t-stat	Degree of freedom	Standard Deviation
Gold	6.0890 6.3392	97.6146	499	1.4235
Silver	41435 42101	246.0079	1071	5.5590e+03
Lead	123.9365 125.3158	354.5992	1070	11.5018
Aluminium	109.8724 110.8075	763.1114	957	7.3745
Copper	68.865 71.096	123.1002	993	17.9232

Parameters discussion

Series of experiments have been performed on tuning different parameter and the optimal value is described in this section. We have used two technical indicators and B and $W\%$, where window size, weight factor, and number of standard deviation values are 20, 0 and 2 respectively. Number of period for indicators $W\%$ is set to be 14. Numbers of neurons set at different layers of BPNN are 11, 15 and 1. There is a huge range for tuning PSO parameter. The population size is set to 30, and the values 100, 2, 2, 0.8, 0.005, initialized to 0 and 50 are taken for $Max\_iteration$, $C1$, $C2$, Dt, $Delm$, P_{best}, G_{best}, respectively. The Inertia weight w is set to 2.

Performance measures

This work includes the specificity, sensitivity and various probabilistic predictive performance measures such as; BCR (Balanced Classification Rate), BER (Balanced error rate), Recall, F-Measure, JI (Jaccard Index), ARI (Adjusted Random Index), NPV (Negative Predicted Value), FPR (False Positive Rate), FNR (False Negative Rate), FDR (False Discovery Rate), F-Score, G-Mean (Geometric Mean), MCC (Mathew's Correlation Coefficient), J-statistics (Youden's J-Statistics), Accuracy (ACC). The value of each measure should lie between [0-1], where 0 represents the lower classification ability and 1 represents the high classification ability.

Result analysis

In this work, the best optimum configuration of the PSO-BPNN model was obtained through MSE, MAPE, and MAD standard error metrics to measure the prediction accuracy. With performance metrics, accuracy on five metal dataset's stock indices has been found as 93.750. The comparison between original and predicted values forclosing price and volatility are shown in Figs 2 and 3. For both cases, the average value of deflecting predicted part from the original found in proposed PSO-BPNN shows smaller and very close to the original value. It indicates that proposed PSO-BPNN is becoming superior in comparison with other algorithms over all datasets.

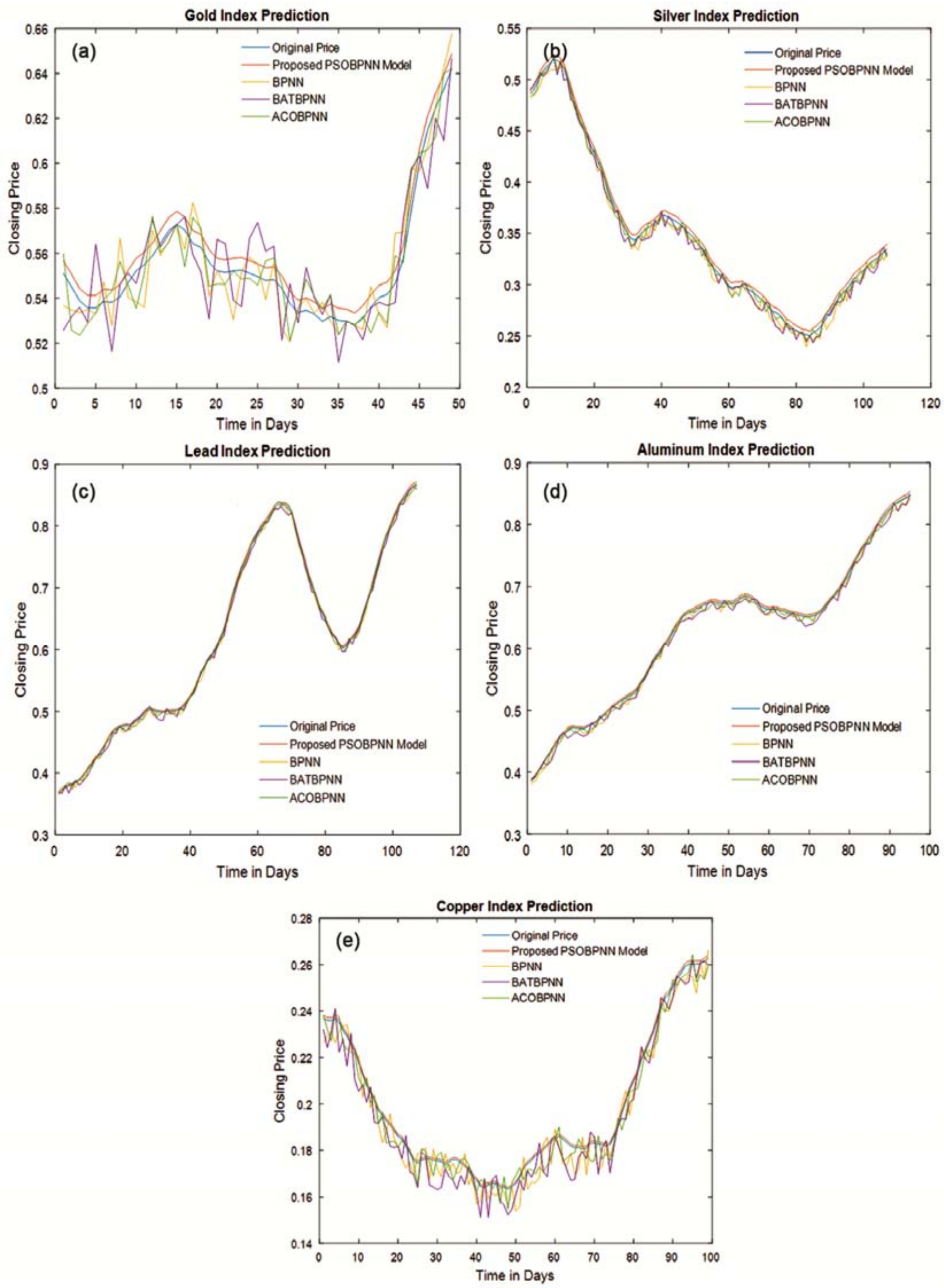

Fig. 2 — Comparison graph for original Closing Index vs. Proposed model output on (a) Gold, (b) Silver, (c) Lead, (d) Aluminium and (e) Copper testing datasets

Data given in Tables 2 and 3 prove that the forecasting decision of closing price and volatility respectively using PSO over other optimization techniques. The corresponding accuracy predicted for closing index and volatility for five datasets are found to be ~ 5 to 10% and ~ 2 to 15% more in comparison with other optimization techniques.

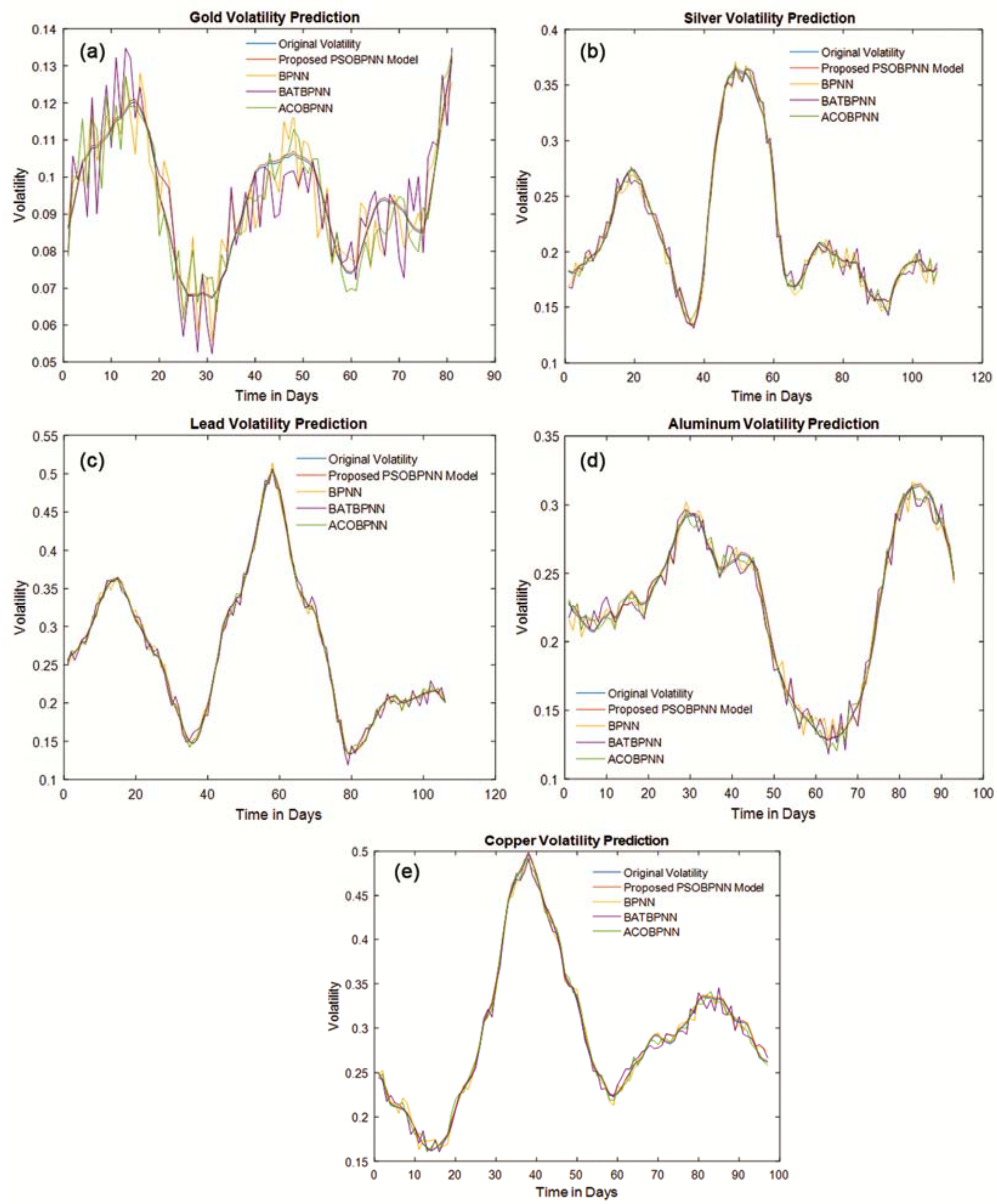

Fig. 3 — Comparison graph for volatility prediction on (a) Gold, (b) Silver, (c) Lead, (d) Aluminium and (e) Copper testing datasets

Table 2 — Comparison of overall prediction accuracy of PSO-BPNN with BPNN, BAT-BPNN, ACO-BPNN network for closing price index

Datasets	Comparison level (%)			Comparison level (%)			Comparison level (%)		
	PSO-BPNN	BPNN	Improvement in (%)	PSO-BPNN	BAT-BPNN	Improvement in (%)	PSO-BPNN	ACO-BPNN	Improvement in (%)
Gold	93.750	75.51	18.24	93.750	86.221	7.529	93.750	81.237	12.513
Silver	91.489	79.43	12.059	91.489	83.412	8.077	91.489	80.113	11.376
Lead	97.826	85.04	12.786	97.826	88.891	8.935	97.826	87.145	10.681
Aluminium	97.778	91.48	6.298	97.778	91.482	6.296	97.778	92.294	5.484
Copper	97.917	84.84	13.077	97.917	86.172	11.745	97.917	89.461	8.456

Table 3 — Comparison of overall prediction accuracy of PSO-BPNN with BPNN, BAT-BPNN, ACO-BPNN network for volatility

Datasets	Comparison level (%)			Comparison level (%)			Comparison level (%)		
	PSO-BPNN	BPNN	Improvement in (%)	PSO-BPNN	BAT-BPNN	Improvement in (%)	PSO-BPNN	ACO-BPNN	Improvement in (%)
Gold	98.76	81.24	17.52	98.76	88.54	10.22	98.76	86.45	12.31
Silver	99.06	82.16	16.9	99.06	89.68	9.38	99.06	84.52	14.54
Lead	99.05	88.68	10.37	99.05	93.18	5.87	99.05	90.28	8.77
Aluminium	98.92	87.14	11.78	98.92	95.32	3.6	98.92	96.48	2.44
Copper	98.96	86.37	12.59	98.96	89.62	9.34	98.96	96.33	2.63

Conclusions and Future Scope

This study proposes a forecasting method based on BPNNs trained by PSO to forecast the commodities' closing price and volatility. In order to score the objective, all three optimized algorithms have been implemented in the context of the five metal markets. From the experimental result analysis it has been found that the proposed PSO-BPNN produces promising output while comparing with BAT-BPNN and ACO-BPNN. For future research, it might use a combined qualitative-quantitative approach, including higher number of influencing factors, to study the different domains or other areas.

Competing Interest

The authors declare that they have no competing interests.

References

1 Nguyen D K & Walther T, Modeling and forecasting commodity market volatility with long-term economic and financial variables, *J Forecast*, **39** (2020) 126–142.

2 Dehghani H & Bogdanovic D, Copper price estimation using bat algorithm, *Resour Policy*, **55** (2018) 55–61.

3 Zhang H, Gu B, Mu J, Ruan P & Li D, Wheat hardness prediction research based on NIR hyperspectral analysis combined with ant colony optimization algorithm, *Procedia eng*, **174** (2017) 648–656.

4 Reddy S S, Bat algorithm-based back propagation approach for short-term load forecasting considering weather factors, *Electr Eng*, **100** (2017) 1–7.

5 Kavousi-Fard A & Su W, A combined prognostic model based on machine learning for tidal current prediction, *IEEE Trans Geosci Remote Sens*, **55** (2017) 3108–3114.

6 Kavousi-Fard A, Khosravi A & Nahavandi S, A new fuzzy-based combined prediction interval for wind power forecasting, *IEEE Trans Power Syst*, **31** (2016) 18–26.

7 Hu Y, Li J, Hong M, Ren J, Lin R, Liu Y & Man Y, Short term electric load forecasting model and its verification for process industrial enterprises based on hybrid GA-PSO-BPNN algorithm—A case study of paper making process, *Energy*, **170** (2019) 1215–1227.

8 Zhang J L, Zhang Y J & Zhang L, A novel hybrid method for crude oil price forecasting, *Energy Econ*, **49** (2015) 649–659.

9 Abdollahi H & Ebrahimi S B, A new hybrid model for forecasting Brent crude oil price, *Energy*, **200** (2020) Article ID 117520.https://doi.org/10.1016/j.energy.2020.117520.

10 Chang J F & Lee Y J, Performance comparison between Genetic Algorithm and Particle Swarm Optimization based on back propagation network in forecasting exchange rate, *J Stat Manag Syst*, **13** (2010) 805–821.

11 Salman N, Lawi A & Syarif S, Artificial neural network back propagation with particle swarm optimization for crude palm oil price prediction, *J Phy Conf Ser*, **1114** (2018, November) 012088.doi :10.1088/1742-6596/1114/1/012088

12 Rehman M Z & Nawi N M, Studying the effect of adaptive momentum in improving the accuracy of gradient descent back propagation algorithm on classification problems, *Int J Mod Phy Conf Ser*, 432–439 2012.

13 Lahmiri S, Interest rate next-day variation prediction based on hybrid feed forward neural network, particle swarm optimization, and multi resolution techniques, *Physica A: Stat Mech Appl*, **444** (2016) 388–396.