Songklanakarin J. Sci. Technol.
43 (4), 1160-1168, Jul. - Aug. 2021

http://www.sjst.psu.ac.th

Original Article

A hybridization of feedforward neural network and differential evolution to forecast fertilizer consumption emphasizing on selecting optimal architecture

Thoranin Sujjaviriyasup*

Department of Logistics Engineering, School of Engineering,
University of the Thai Chamber of Commerce, Din Daeng, Bangkok, 10400 Thailand

Received: 19 January 2020; Revised: 24 May 2020; Accepted: 2 September 2020

Abstract

A fertilizer marketing or producing sector has played an important role in agricultural productivity and also food security around the world for many years. However, the demand of fertilizer consumption is uncertain and difficult to be forecasted by using simple approaches. Therefore, an accuracy of future demand concerning fertilizer is very interesting task to support decision making. In this research, a hybrid model of feedforward neural networks and differential evolution emphasizing on architectural evolution is developed to forecast ten datasets of fertilizer consumption and is compared with conventional models based on five accuracy measures. The empirical results indicated that the developed model can provide more accuracy than conventional models at 0.05 significance levels. Furthermore, the capability of the developed model can also provide the highest precision compared with both ARIMA and SVR models. Consequently, the developed model can be a promising tool to predict future demands of fertilizer consumption.

Keywords: time series analysis, combined model, agro - industry, parameter selection, optimization technique

1. Introduction

From an agro-economic point of view, a fertilizer marketing or producing sector has played a significant role in agricultural productivity and also food security all over the world (Cordell, Drangert, & White, 2009; Geman & Eleuterio, 2013; Kopittke *et al.*, 2019; Stewart & Roberts, 2012) for many years. Consequently, a numerous demand of fertilizer has been continuously increasing in each year. According to the high consumption of fertilizer, it is capital intensive and cost sensitive for agriculturists as well (Komarek, *et al.*, 2017). Moreover, the pattern of fertilizer consumption depends heavily on either a variety of agro-economic factors or climate change (Ganesan & Raut, 2012; Mârza, Angelescu, & Tindeche, 2015) in each crop year, which is not stable nor is it easy to access to gather information. For example, actual area cultivated in each crop is not constant due to a marketing situation. Besides, an extent of available land is difficult to be estimated accurately and timely. Even flood or drought situations which agriculturists often experience in each crop year.

For agriculturists, the forecast is used for a guideline by which the agriculturists evaluate operational effectiveness (Ganesan & Raut, 2012; Mishra, Sahu, & Uday, 2014). For instance, the demand forecast of fertilizer can support a proper purchasing order as close to seasonal demand as possible. Moreover, the demand forecast can also support an inventory plan to maintain working capital within control. Meanwhile, the demand forecast at nation level can support a government to monitor the overall supply – demand balance in order to estimate demand and ensure adequate overall fertilizer availability. Likewise, the demand forecast can provide useful information to conserve foreign exchange and also reduce an economic burden of excessive inventories. Therefore, the future fertilizer consumption is required to realize before making a critical decision on appropriate planning. However,

*Corresponding author
Email address: thoranin_suj@utcc.ac.th

the demand pattern is fluctuating and difficult to be estimated accurately by using simple approaches. Subsequently, several forecasting approaches are developed and proposed to predict the future demand. Thus, the accuracy of future demand is a very interesting issue to support the decision making regarding supply chain of fertilizer in present. One of several popular forecasting techniques is time series analysis (Farajian, Moghaddasi, & Hosseini, 2018; Kyriazi, Thomakos, & Guerard, 2019; Nagy, Fehér & Tamás, 2018) that is formulated from previous observations and is lower dimensional data than regression analysis. Consequently, the time series analysis is a more convenient approach to forecast the future demand of fertilizer in real-world problems unless many influence factors are available and stable for regression analysis.

In a field of statistical forecasting techniques concerning time series analysis, an autoregressive integrated moving average (ARIMA) model is a well-known approach and has dominated in linear forecasting problems in many years (Babai, Ali, Boylan, & Syntetos, 2013; He & Tao, 2018; Sen, Roy, & Pal, 2016). Additionally, the ARIMA model is always employed as benchmark to compare with other forecasting models in M3 competitions. Even though the ARIMA model is successful in linear forecasting problems, it may not be appropriate and is not flexible for all circumstances including many nonlinear forecasting problems in real-world. Nevertheless, an analysis of time series is still difficult to identify whether a time series is generated from linear or nonlinear underlying process at present. Consequently, the ARIMA model is still used in many fields of science including agricultural sector in recent years (Ganesan & Raut, 2012; Ohyver & Pudjihastuti, 2018).

With regard to solve nonlinear forecasting problems, many supervised machine learning models are developed and proposed during recent years. One of the well-known models is support vector regression (SVR) model, which achieves an optimum network structure and always provides globally optimal solution based on convex optimization problem. Consequently, the SVR model has more attractive for nonlinear forecasting problems in many fields of science (Al-Musaylh, Deo, Adamowski, & Li, 2018; Chen, *et al.*, 2017). Moreover, the SVR model is also developed and applied to many literatures of agricultural sector at present (Sujjaviriyasup, 2018; Su, Xu, & Yan, 2017).

Although the SVR model has overwhelmingly attractive in nonlinear forecasting problems, the forecasting performance of SVR models depends heavily upon a suitable hyper-parameter selection that is a major concern to improve forecasting performance of SVR model. Therefore, there is a risk of using inappropriate parameters of SVR model unless the optimization techniques are adopted. On the other hand, another well-known model is artificial neural networks (ANNs) (Co & Boosarawongse, 2007; Constantino, Fernandes, & Teixeira, 2016; Li, Xu, & Li, 2010; Zou, Xia, Yang, & Wang, 2007) that were intended for simulating network system of biological neurons based on empirical risk minimization principle. Consequently, the ANN models tend to encounter a problem of overfitting rather than SVR models. However, since these two well-known models have their own advantages and disadvantages, it is difficult to decide which one has superior capability in forecasting (Ahmad *et al.*, 2014). Thus, ANN models have been still applied in forecasting and are also more widely implemented in many fields of science including time series analysis. One of the most popular ANN models is feedforward neural network (FFNN) model with one hidden layer, which is also successful in time series analysis (Lolli *et al.*, 2017; Qiao, Li, Han, & Wang, 2017; Wu & Wang, 2012).

The FFNN model consists of three layers that are input layer, hidden layer, and output layer. Nevertheless, the performance of FFNN model depends heavily on either architecture or learning algorithm. Pertaining to investigate suitable FFNN model, the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm is employed to adjust weights of FFNN model to improve forecasting accuracy. In addition, the BFGS algorithm is quite efficient to find the suitable weights of FFNN model. On the other hand, the architectures of FFNN model do not depend on any algorithm to formulate FFNN structure. Consequently, many methods are proposed to search the most appropriate architecture of FFNN models in recent years.

One of the most useful methods with regard to appropriate architecture of FFNN models is metaheuristics (Han, Jiang, Ling, & Su, 2019; Ojha, Abraham, & Snášel, 2017) that may provide a sufficiently good solution for optimization problems. In other words, the metaheuristics can provide a globally optimal architecture of FFNN model within given search space and also may be usable for a variety of problems. Furthermore, the metaheuristics are often able to find good architecture of FFNN model with less computational effort than other optimization algorithms.

Differential evolution (DE) algorithm is one of several useful search methods (Arce, Zamora, Sossa & Barrón, 2018; Han, Li, Wu, Zhu, & Song, 2019; Sujjaviriyasup, 2019; Yang, Chen, Wang, Li, & Li, 2016), which was intended from natural inspiration with regard to evolutionary. Consequently, the DE is a population based stochastic search algorithm that consists of crossover, mutation, and selection operations on a population to minimize an objective function. Moreover, the DE algorithm is considered the most recent evolutionary algorithms.

In this paper, a hybrid forecasting model is developed and proposed to predict ten datasets of annual fertilizer consumption. The proposed model exploits FFNN models to formulate complex prediction function while the DE algorithm is employed to search the appropriate architecture of FFNN models within a given search space. The motivation of proposed model aims at adopting DE algorithm to reduce a risk of using improper architecture concerning the FFNN models and improve accuracy of forecasting as well. In order to evaluate forecasting performance of the proposed model is compared with ARIMA and SVR based on five accuracy measures that are mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), symmetric mean absolute percentage error (sMAPE), and root mean square percentage error (RMSPE). Concerning these accuracy measures, MAPE is recommended to compare across datasets and may still be preferred for reasons of simplicity to explain regarding forecast accuracy. Moreover, an analysis of variance and a post hoc test based on MAPE are exploited to identify the significant difference among three forecasting models.

2. Materials and Methods

2.1 Datasets of fertilizer consumption

All datasets of fertilizer consumption used in this research are annual time series datasets of each nation around the world, which are obtained from online datasets of the World Bank (The World Bank, 2019) as shown in Figure 1.

2.2 Methodologies

In this sector, all forecasting models are presented to describe mathematical formulation. However, the well-known models are demonstrated in brief. Meanwhile, the developed forecasting model is described more details in order to focus on complex procedure of the developed model.

2.2.1 Autoregressive integrated moving average model

The autoregressive integrated moving average has good performance for many linear problems of time series analysis and is also generalization of autoregressive moving average with regard to non-stationary time series data. The ARIMA(*p*, *d*, *q*) model with mean μ is general form of the autoregressive integrated moving average, which has mathematical formula as Equation (1).

$$\left(1-\sum_{i=1}^{p}\varphi_i B^i\right)(1-B)^d (y_t-\mu)=\left(1-\sum_{j=1}^{q}\theta_j B^j\right)\varepsilon_t \quad (1)$$

where y_t and ε_t are time series data and random error at time period *t*, respectively. Meanwhile, φ and θ are model parameters; *p* and *q* are referred to as orders of autoregressive integrated moving average. The *B* and *d* are the backward shift operator and degree of differencing.

The most suitable ARIMA model is selected from the lowest Akaike Information Criterion with a correction for finite sample sizes (AICc). The mathematical formulation of AICc is described as Equation (2).

$$\text{AICc}=2k-2\ln(L)+\frac{2k(k+1)}{n-k-1} \quad (2)$$

where *L* is the maximum value of likelihood function for the ARIMA model. The *n* and *k* are sample size of time series data and parameters of ARIMA model, respectively. The automated function, namely auto.arima, is used in this article (Hyndman and Athanasopoulos, 2018; Hyndman, *et. al.*, 2015). However, the most suitable ARIMA model relies on the change of previous observation update. Therefore, the ARIMA is used to stand for ARIMA(*p*,*d*,*q*) with mean μ in this article.

2.2.2 Support vector regression model

The support vector regression is a supervised machine learning model, which is extended from support vector machine to address regression problems. The general formulation of $\varepsilon-$SVR model with regard to either linear or nonlinear regression tasks is presented as Equation (3).

$$f(x_i)=\sum_{i=1}^{T}\left(\alpha_i-\alpha_i^*\right)K(x,x_i)+b \quad (3)$$

where α_i and α_i^* are the so-called Lagrange multipliers, *b* is a scalar threshold, $(\cdot,\cdot)$ denote vector inner product, and $K(x,x_i)$ is kernel functions. In general, four types of kernel function satisfying Mercer's condition can be used as the kernel functions of the SVR that are described as follows in Equations (4) – (7).

Linear: $K(x,x_i)=x^T x_i$ (4)

Polynomial: $K(x,x_i)=\left(\gamma x^T x_i+r\right)^p$ (5)

Radial basis: $K(x,x_i)=\exp\left(-\gamma\|x-x_i\|^2\right)$ (6)

Sigmoid: $K(x,x_i)=\tanh\left(\gamma x^T x_i+r\right)$ (7)

where γ, *r*, and *p* are kernel parameters.

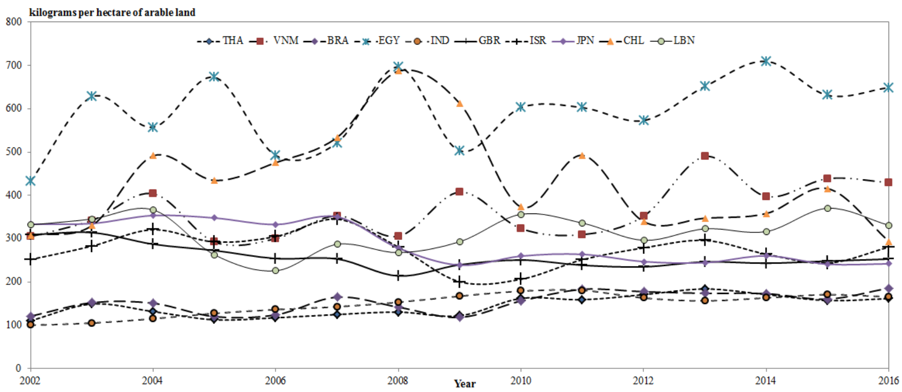

Figure 1. The annual datasets of fertilizer consumption

For time series analysis with SVR model in this article, the time series dataset is rearranged into *m* columns. The first *m* – 1 columns of the matrix of time series dataset are exploited as input data. Meanwhile, the last column of the matrix of time series dataset is employed as target data. The SVR(*m*) is referred to as SVR model with *m* columns to train and formulate model. In addition, the suitable SVR model is formulated by using grid search to find the proper parameters of SVR model. The data preparation of time series dataset before applying the SVR model is demonstrated in Figure 2.

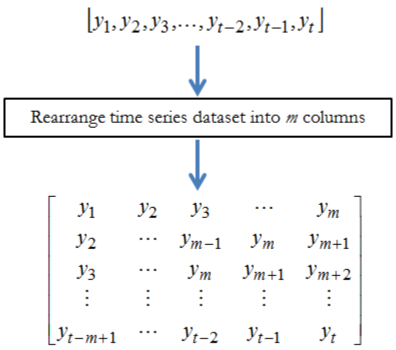

Figure 2. The data preparation of time series dataset

2.2.3 Differential evolution optimization

The differential evolution is one of the most powerful stochastic real-parameter optimization algorithms (Storn & Price, 1997), which operates similar computational steps of standard evolutionary algorithm. The DE algorithm is illustrated as follows:

1. The objective function and all parameters of DE algorithm are defined.
2. All agents x are generated with random positions in the search space.
3. Until a termination criterion is met repeat the follows:
   - Three members of the population are selected to formulate an initial mutant parameter vector.
   - A trial vector is generated by using crossover operator.
   - The differential evolution adopts a greedy selection operator.
4. The highest fitness or lowest value of objective function is selected from the population and return it as the best found candidate solution.

2.2.4 Hybridization of feedforward neural network and differential evolution

The hybrid of feedforward neural network and differential evolution is developed to forecast fertilizer consumption. The motivation of developed model is to formulate a complex forecasting model by using FFNN model to establish nonlinear prediction function while the DE algorithm is employed to select the optimal architecture of FFNN model emphasizing on neurons in hidden layer. The algorithm of developed model is presented as follows:

For *m* is equal to 2 **to** a termination criterion **do**

• The DE algorithm generates initial parameters based on dimensional features of the FFNN model as presented in Equation (8).

$$\theta_{i,G} = \lfloor \theta_{1,i,G}, \theta_{2,i,G}, \theta_{3,i,G}, \ldots, \theta_{d,i,G} \rfloor \quad i = 1,2,3,\ldots, N \tag{8}$$

where θ is a vector of nodes in hidden layer, d is dimensional nodes of the FFNN model, G is the number of generation, and N is the size of population.

• Until a termination criterion is met repeat the follows:

For each agent θ in the population **does**:

For *t* is equal to 67% of the previous observations **to** the observation before the last observation **do**

• Rearrange the previous observations into *m* columns of the previous observations.

$$\begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_m \\ y_2 & y_3 & y_4 & \cdots & y_{m+1} \\ y_3 & y_4 & y_5 & \cdots & y_{m+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t-m+1} & \cdots & y_{t-2} & y_{t-1} & y_t \end{bmatrix}$$

• For the first *m* – 1 columns of the matrix of the previous observations are used as input data, while the last column of the matrix of the previous observations is adopted as target data.

• Until a termination criterion is met repeat the follows:

• Convert θ from real number to integer number and utilize θ as node numbers in hidden layer of the FFNN model to generate weights corresponding to networks connect between input nodes and hidden nodes as Equation (9)

$$D_j = g\left(w_{0j} + \sum_{i=1}^{n} w_{ij} y_{t-(i-1)} \right) \tag{9}$$

where $g(\bullet)$ is sigmoid function to transform time series dataset into nonlinear space, w_{ij} is the weight between node *i* in input layer and node *j* in hidden layer, w_{0j} is bias of each node in hidden layer, *n* is the number of input node, and D_j is node *j* of hidden layer when number *j* is equal to θ.

• Exploit the D_j as new input dataset to formulate linear function and predict future value as Equation (10)

$$y_{t+1} = h\left(\beta_0 + \sum_{j=1}^{\theta} \beta_j D_j\right) \tag{10}$$

where $h(\bullet)$ is linear function to predict future value, β_j is weight between node *j* in hidden layer and node of output layer, β_0 is bias of output node, and D_j is node *j* of hidden layer.

• Adjust weights of FFNN model by using BFGS algorithm.

• Forecast future value.

• Calculate MAPE.

End

• An initial mutant parameter vector $v_{i,G}$ is created from selecting three members of the population, $\theta_{r_0,G}$, $\theta_{r_1,G}$ and $\theta_{r_2,G}$ at random as equation (11).

$$v_{i,G} = \theta_{i,G} + (\theta_{best,G} - \theta_{i,G}) + \theta_{r_0,G} + 0.8 \times (\theta_{r_1,G} - \theta_{r_2,G}) \tag{11}$$

where $\theta_{i,G}$ and $\theta_{best,G}$ are the *i*-th vector of the population at the current generation and the best individual vector with the best fitness, respectively. *G* is the number of generations; r_0, r_1, and r_2 are randomly chosen numbers within the population size; and $i = 1,2,3,...,N$

• A trial vector $u_{i,G}$ is generated by using the crossover operator as Equation (12)

$$u_{i,G} = \begin{cases} v_{j,i,G} & if \;\; rand_{j,i} \le 0.5 \text{ or } j = I_{rand} \\ \theta_{j,i,G} & otherwise \end{cases} \tag{12}$$

where $i = 1,2,3,...,N$; $j = 1,2,3,...,d$, $rand_{j,i} \sim U[0,1]$, I_{rand} is a random integer from $[1,2,...,d]$. *G* is the number of generations.

• A greedy selection operator is adopted in the differential evolution process as Equation (13).

$$\theta_{i,G+1} = \begin{cases} u_{i,G} & if \;\; f(u_{i,G}) \le f(\theta_{i,G}) \\ \theta_{i,G} & otherwise \end{cases} \tag{13}$$

where $f(u_{i,G})$ is the MAPE of the trial vector and $f(\theta_{i,G})$ is equal to MAPE of the target vector. *G* is the number of generations; and $i = 1,2,3,...,N$

End

• Choose the agent from the population that has the lowest MAPE and return it as the best found parameters of FFNN model.

End

Regarding time series analysis with the developed model, FFNN(*m*, θ)-DE is referred to as the FFNN model with *m* columns and θ nodes in one hidden layer that is found by using differential evolutionary algorithm.

2.3 Cross-validation

In general, there is not any forecasting model that has good performance in many situations. Consequently, all forecasting models have to be evaluated their forecast accuracy by using ten time series datasets of fertilizer consumption in each country, which are obtained from the World Bank. Each dataset of the fertilizer consumption is partitioned to two subsets as training dataset and test dataset. For forecast modeling, the training dataset should be sufficient to investigate the most suitable model and also provide the forecast accuracy. Meanwhile, the test dataset should be enough to validate the forecasting performance. Nonetheless, the limitation of time series length regarding fertilizer consumption exists. Therefore, the datasets of fertilizer consumption are separated into 67% and 33% for training dataset and test dataset, respectively. Regarding training dataset, approximately 67% of each time series dataset of fertilizer consumption is exploited to form the fitted model and to predict one step ahead. The rest of each time series dataset of fertilizer consumption is used to evaluate those forecasting models as hold – out set. After the actual data is presented, then it is gathered into training dataset to model and predict one step ahead until the last data of hold – out set.

In order to indicate accuracy of forecast, five existing measures of accuracy are used to evaluate all forecasting models, which are the most commonly used measures both scale – dependent measure and scale – independent measure. For scale – dependent measure, MAE and RMSE measures is often recommended to evaluate forecast accuracy. On the other hand, MAPE, sMAPE, and RMSPE measures are commonly used to evaluate forecast accuracy across different datasets due to advantage of being scale – independent. All mathematical expression of the accuracy measures are demonstrated in Equations (14) – (18).

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \tag{14}$$

$$\mathrm{RMSE} = \sqrt{\left(\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}\right)} \tag{15}$$

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i| / y_i}{n} \times 100 \tag{16}$$

$$\mathrm{sMAPE} = \frac{\sum_{i=1}^{n} 2 \times |y_i - \hat{y}_i| / (y_i + \hat{y}_i)}{n} \times 100 \tag{17}$$

$$\mathrm{RMSPE} = \sqrt{\frac{\sum_{i=1}^{n} ((y_i - \hat{y}_i)/y_i \times 100)^2}{n}} \tag{18}$$

3. Results and Discussion

For evaluating superior performance of forecasting models in this article, the cross-validation and five accuracy measures are exploited to investigate the most suitable model with the lowest errors in many situations. In order to demonstrate forecasting performance in each period, the forecasting models regarding fertilizer consumption of Thailand are selected and presented in Figure 3. The summary of forecasting models based on five accuracy measures is presented in Table 1.

In table 1, the empirical results indicated that the SVR model as supervised machine learning for nonlinear problems outperforms ARIMA model as statistical model for many situations. However, the developed model provides the lowest errors in all situations. In other words, the proposed model can provide more accuracy than other forecasting models. This evidence indicated that the developed model has superior capability of forecasting than both ARIMA and SVR models. In addition, a box plot is employed to describe graphical examination of MAPE concerning all forecasting models as shown in Figure 4.

As the results of box plot, the developed model demonstrated that it provides both the lowest mean and median of MAPE compared with both ARIMA and SVR models. Moreover, the dispersion of the developed model is less than other candidate models.

In order to identify significant difference among performances of forecasting model, an analysis of variance based on nonparametric test is adopted to analyze due to its ability of distribution-free tests. Moreover, the given results of normality test indicated that MAPE datasets of each forecasting model does not come from a normally distributed population at 0.05 significance levels. This evidence can support to use nonparametric test dealing with this problem rather than parametric test that based on normally distributed population. One of several nonparametric tests is Friedman test that is firstly exploited for one-way repeated measures analysis of variance by ranks. The null hypothesis of Friedman test is there are no differences among forecasting

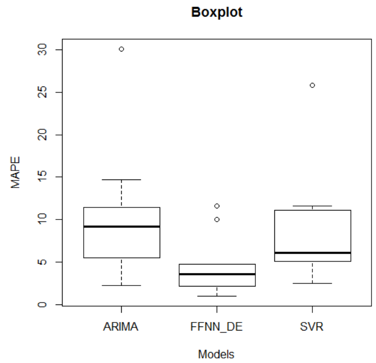

Figure 4. The graphical examination of box plot based on MAPE

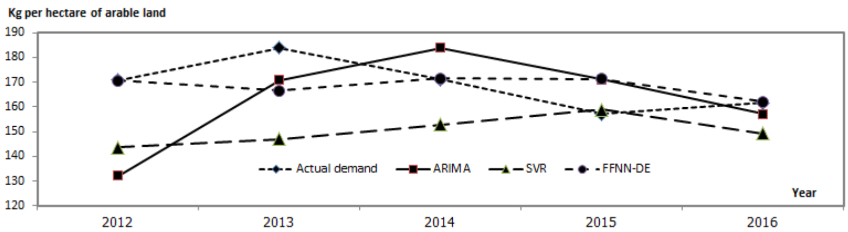

Figure 3. The forecasting models concerning fertilizer consumption of Thailand

Table 1. The summary of all forecasting models based on five accuracy measures

Dataset	Model	MAE	RMSE	MAPE	sMAPE	RMSPE
	ARIMA	16.656	20.293	9.81	10.32	11.92
Thailand	SVR(7)	19.305	22.766	11.09	11.99	12.88
	FFNN(7,150)-DE	6.553	10.035	3.84	3.85	5.84
	ARIMA	65.611	82.869	14.67	16.48	17.76
Vietnam	SVR(8)	53.199	71.695	11.62	12.92	15.08
	FFNN(6,91)-DE	45.807	64.419	10.01	11.04	13.60
	ARIMA	20.113	23.397	11.41	12.24	13.20
Brazil	SVR(6)	17.784	20.350	9.99	10.68	11.38
	FFNN(7,81)-DE	8.483	10.507	4.74	4.88	5.80
	ARIMA	76.409	92.413	11.45	12.45	13.43
Egypt, Arab Rep.	SVR(7)	44.953	56.166	6.75	7.07	8.18
	FFNN(3,98)-DE	28.151	34.935	4.21	4.34	5.12
	ARIMA	14.023	15.537	8.60	8.34	9.55
India	SVR(3)	4.775	6.268	2.89	2.82	3.78
	FFNN(6,65)-DE	1.937	2.253	1.19	1.19	1.39
	ARIMA	5.601	6.376	2.28	2.30	2.59
United Kingdom	SVR(2)	6.094	6.683	2.49	2.50	2.74
	FFNN(3,68)-DE	2.545	3.343	1.03	1.03	1.36
	ARIMA	15.252	17.310	5.52	5.64	6.17
Israel	SVR(3)	13.447	15.867	5.12	4.96	6.27
	FFNN(3,10)-DE	9.123	14.815	3.42	3.29	5.59
	ARIMA	11.096	13.476	4.47	4.40	5.44
Japan	SVR(4)	12.492	15.975	5.07	4.89	6.49
	FFNN(8,50)-DE	5.313	6.174	2.17	2.14	2.53
	ARIMA	100.588	107.703	30.07	25.55	32.80
Chile	SVR(8)	85.774	99.953	25.79	22.02	30.55
	FFNN(6,70)-DE	40.987	56.367	11.63	10.82	16.31
	ARIMA	26.788	33.407	7.82	8.04	9.42
Lebanon	SVR(8)	18.218	26.744	5.41	5.51	7.76
	FFNN(6,50)-DE	10.865	20.122	3.03	3.19	5.45

performances. On the other hand, the alternative hypothesis is that at least one of forecasting performances is significantly different from all of the others. The summary of Friedman test based on MAPE is illustrated in Figure 5.

With regard to p-value of Friedman test, the null hypothesis is rejected at 0.05 significance levels. Consequently, this evidence is sufficient to support that at least one of forecasting performances is significantly different from all of the others. Subsequently, the post hoc test is employed to identify significant difference between the forecast performances of these models. The summary of pairwise comparisons using Conover's test is demonstrated in

The results of pairwise comparison in Figure 6, the SVR model as machine learning model outperforms the ARIMA model as statistical model at 0.05 significance levels. Moreover, the developed model provides significantly lower error than both ARIMA and SVR models at 0.05 significance levels. All evidences are sufficient to support that the developed model, which is formulated from FFNN model and differential evolution, can provide more accuracy than both ARIMA and SVR models at 0.05 significance levels. In other words, the hybrid forecasting model outperforms these single conventional models at 0.05 significance levels.

```
        Friedman rank sum test

data:  data$MAPE, data$Model and data$Block
Friedman chi-squared = 15.8, df = 2, p-value = 0.0003707
```

Figure 5. The summary of Friedman test based on MAPE

```
        Pairwise comparisons using Conover's test for a two-way
                      balanced complete block design

data:  data$MAPE , data$Model and data$Block

        ARIMA   FFNN_DE
FFNN_DE 6.4e-09 -
SVR     0.027   3.4e-07

P value adjustment method: none
```

Figure 6. The summary of Conover test based on MAPE

4. Conclusions

According to all investigations in many situations, they indicated that the developed model has superior ability with the most accuracy compared with candidate models at 0.05 significance levels and the highest precision dealing with fertilizer consumption. Furthermore, it can support to infer that the optimal structure of FFNN model with differential evolution can reduce a risk of using improper structure of FFNN model and can provide more accuracy than conventional models. Hence, the developed model reveals that it is able to be a promising approach with regard to predict future fertilizer consumption. In addition, the developed model may provide meaningful guidelines for policy makers to make critical decision on efficient planning.

References

Ahmad, A. S., Hassan, M. Y., Abdullah, M. P., Rahman, H. A., Hussin, F., Abdullah, H., & Saidur, R. (2014). A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews, 33*, 102-109.

Al-Musaylh, M. S., Deo, R. C., Adamowski, J. F., & Li, Y. (2018). Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. *Advanced Engineering Informatics, 35*, 1-16.

Arce, F., Zamora, E., Sossa, H., & Barrón, R. (2018). Differential evolution training algorithm for dendrite morphological neural networks. *Applied Soft Computing, 68*, 303-313.

Babai, M. Z., Ali, M. M., Boylan, J. E., & Syntetos, A. A. (2013). Forecasting and inventory performance in a two-stage supply chain with ARIMA (0, 1, 1) demand: Theory and empirical analysis. *International Journal of Production Economics, 143*(2), 463-471.

Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., . . . Wang, K. (2017). Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Applied Energy, 195*, 659-670.

Co, H. C., & Boosarawongse, R. (2007). Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks. *Computers and Industrial Engineering, 53*(4), 610-627.

Constantino, H. A., Fernandes, P. O., & Teixeira, J. P. (2016). Tourism demand modelling and forecasting with artificial neural network models: The Mozambique case study. *Tékhne, 14*(2), 113-124.

Cordell, D., Drangert, J. O., & White, S. (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change, 19*(2), 292-305.

Farajian, L., Moghaddasi, R., & Hosseini, S. (2018). Agricultural energy demand modeling in Iran: Approaching to a more sustainable situation. *Energy Reports, 4*, 260-265.

Ganesan, V. K., & Raut, S. (2012). Demand forecasting for fertilizers—A tactical planning framework for industrial use. 126 Agro-Informatics and Precision Agriculture 2012, 123-129.

Geman, H., & Eleuterio, P. V. (2013). Investing in fertilizer–mining companies in times of food scarcity. *Resources Policy, 38*(4), 470-480.

Han, F., Jiang, J., Ling, Q. H., & Su, B. Y. (2019). A survey on metaheuristic optimization for random single-hidden layer feedforward neural network. *Neurocomputing, 335*, 261-273.

Han, J. W., Li, Q. X., Wu, H. R., Zhu, H. J., & Song, Y. L. (2019). Prediction of cooling efficiency of forced-air precooling systems based on optimized differential evolution and improved BP neural network. *Applied Soft Computing, 84*, 105733.

He, Z., & Tao, H. (2018). Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study. *International Journal of Infectious Diseases, 74*, 61-70.

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.

Hyndman, R. J., Athanasopoulos, G., Razbash, S., Schmidt, D., Zhou, Z., Khan, Y., . . . & Wang, E. (2015). Forecast: Forecasting functions for time series and linear models. R package version, 6(6), 7.

Komarek, A. M., Drogue, S., Chenoune, R., Hawkins, J., Msangi, S., Belhouchette, H., & Flichman, G. (2017). Agricultural household effects of fertilizer price changes for smallholder farmers in central Malawi. *Agricultural Systems, 154*, 168-178.

Kopittke, P. M., Menzies, N. W., Wang, P., McKenna, B. A., & Lombi, E. (2019). Soil and the intensification of agriculture for global food security. *Environment International, 132*, 105078.

Kyriazi, F., Thomakos, D. D., & Guerard, J. B. (2019). Adaptive learning forecasting, with applications in forecasting agricultural prices. *International Journal of Forecasting, 35*(4), 1356-1369.

Li, G. Q., Xu, S. W., & Li, Z. M. (2010). Short-term price forecasting for agro-products using artificial neural networks. *Agriculture and Agricultural Science Procedia, 1*, 278-287.

Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T., & Gucci, S. (2017). Single-hidden layer neural networks for forecasting intermittent demand. *International Journal of Production Economics, 183*, 116-128.

Mârza, B., Angelescu, C., & Tindeche, C. (2015). Agricultural insurances and food security. The new climate change challenges. *Procedia Economics and Finance, 27*, 594-599.

Mishra, P., Sahu, P. K., & Uday, J. P. S. (2014). ARIMA modeling technique in analyzing and forecasting fertilizer statistics in India. *Trends in Biosciences, 7*(3), 170-176.

Nagy, A., Fehér, J., & Tamás, J. (2018). Wheat and maize yield forecasting for the Tisza river catchment using MODIS NDVI time series and reported crop statistics. *Computers and Electronics in Agriculture, 151*, 41-49.

Ohyver, M., & Pudjihastuti, H. (2018). Arima model for forecasting the price of medium quality rice to anticipate price fluctuations. *Procedia Computer Science, 135*, 707-711.

Ojha, V. K., Abraham, A., & Snášel, V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence, 60*, 97-116.

Qiao, J., Li, S., Han, H., & Wang, D. (2017). An improved algorithm for building self-organizing feedforward neural networks. *Neurocomputing, 262*, 28-40.

Sen, P., Roy, M., & Pal, P. (2016). Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy, 116*, 1031-1038.

Stewart, W. M., & Roberts, T. L. (2012). Food security and the role of fertilizer in supporting it. *Procedia Engineering, 46*, 76-82.

Storn, R., & Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization, 11*(4), 341-359.

Su, Y. X., Xu, H., & Yan, L. J. (2017). Support vector machine-based open crop model (SBOCM): Case of rice production in China. *Saudi Journal of Biological Sciences, 24*(3), 537-547.

Sujjaviriyasup, T. (2018). Predicting prices of agricultural commodities in Thailand using combined approach emphasizing on data pre-processing technique. *Songklanakarin Journal of Science and Technology, 40*(1).

Sujjaviriyasup, T. (2019). Forecasting petroleum consumption using hybrid SVR-DE model emphasizing on optimal parameter selection technique. *Songklanakarin Journal of Science and Technology, 41*(6).

The World Bank. (2019, October 3). Fertilizer consumption. Retrieved from https://data.worldbank.org/indicator/AG.CON.FERT.ZS

Wu, X., & Wang, Y. (2012). Extended and Unscented Kalman filtering based feedforward neural networks for time series prediction. *Applied Mathematical Modelling, 36*(3), 1123-1131.

Yang, Y., Chen, Y., Wang, Y., Li, C., & Li, L. (2016). Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting. *Applied Soft Computing, 49*, 663-675.

Zou, H. F., Xia, G. P., Yang, F. T., & Wang, H. Y. (2007). An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. *Neurocomputing, 70*(16-18), 2913-2923.