Support vector regression and Adaptive neuro fuzzy to measure the Bullwhip effect in supply chain

Edy Fradinata*1, Zurnila Marli Kesuma2, Siti Rusdiana3

1Department of Industrial Engineering, Faculty of Engineering, Syiah Kuala University, Banda Aceh, Indonesia.
2Department of Statistics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Banda Aceh, Indonesia.
3Department of Mathematics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Banda Aceh, Indonesia.

*Email: edinata69@gmail.com

Abstract. Support Vector Regression (SVR) and Adaptive Neuro-Fuzzy (ANFIS) are the advance statistic knowledge. They are employed to construct the forecasting signal models for a tin milk industry. Then the models are used to measure the Bullwhip Effect (BWE) in the supply chain with the beer game methods. This study tried to minimize the BWE with forecasting methods through the comparative for both methods and attempt to prove that a small Mean Square Error (MSE) of the BWE. The data is collected from the price and demand variables, then it is generated in random normal distribution. It informs that in two conclusions. ANFIS is better than SVR in case of MSE comparison. Moreover, the totals BWE of ANFIS and SVR are 209 and 1,237 respectively. The Second conclusion is the better forecast model leads to reduce the amplification across the supply chain.

Key words: SVR, ANFIS, beer game, MSE, BWE

1. Introduction

The supply chain is an integration of information and the physical with involve the coordinating to get a smooth of good flow. Every chain will involve two different activities, physical and information flow. The information will be able to maintain the condition of production planning and to manage inventory in a better way. Managing inventory to keep lower buffers could contribute to the profit for the company. Information from the customer to the retailer helps the management make a better decision [1].

In general, the demand is still difficult to predict because of customer behavior and the environment of communities such as the culture and local or national government regulations. The SVR and ANFIS have become more popular to be used to solve this problem. These methods provide better performance than traditional techniques [2]. SVR is a method that it involves the training points in a small subset to get an output solution with the computer. The loss functions of SVR pretend the optimization in the global minimum area. SVR can improve various interesting features and produce a better performance [3]. ANFIS is one of the AI methods. The ANFIS model contains artificial neural network and fuzzy. This approach was built by Sugeno Fuzzy Inference Model (SFIM), where it built and made the input-output mapping with if-then rules and manipulate them into the system to get the inference of an imprecise model [4].

The objectives of this research are exploring data mining methods of SVR and ANFIS. The second thing is to reduce the BWE with tin milk demand dataset in the supply chain.
2. Literature Review

2.1. Supply Chain and Bullwhip Effect in Tin Milk Industry

The flow of supply chain formed by two components, real physical flow and information flow where they must be managed properly to protect the bullwhip effect in across the supply chain. The supply chain is developed in two different areas, a network of supply chain planning and the execution [5]. BWE is the variability of demand information or distorts from a customer to upstream across the supply chain members [6].

One of the models of variance modeling predicting is called forecasting [7, 8]. The Forecasting model is the most important approach to control and recognize. The control is why the amplification occurs in the member of supply chain in Figure 1. There are five leading causes of the amplification: demand forecast, batch order, price fluctuations, supply shortages, and non-zero lead-time [9]. The wave of bullwhip effect can be seen in Figure 1.

![Bullwhip effect in the supply chain](image)

**Figure 1. Bullwhip effect in the supply chain**

The BWE will occur if the ratio more than one, (BWE > 1). In the case of BWE < 1, it considered the system was well controlled. The formula of a bullwhip effect is presented in the equation 1.

\[
\text{Bullwhip Effect} = \frac{\text{Variance(order)}}{\text{Variance (demand)}} \tag{1}
\]

The retailer realizes the variance of customer demand is named variance demand followed by the wholesaler and distributor, respectively [10]. Massachusetts Institute of Technology (MIT) School developed the beer game simulation for studying the supply chain system on the year of 1960. The role-playing was taking all members of the supply chain integrated each other. The software has three assumptions as follow:
1). Where when the inventory position falls below \( s \), the system places an order \( Q \) (s-Q Policy).
2). Demand follows a Random Normal distribution during the whole 24 weeks.
3). The order Interval (time elapsed between 2 orders) is a fixed at the one-week period.

Using the nominations above now we use relevant formulas as below [11] :

- Base-stock-level is \( (R + L) \cdot \text{AVG} \).
- Safety stock calculated is \( z \cdot \text{STD} \cdot (R+L)^{1/2} \).
- Expected level of inventory is \( r \cdot \text{AVG} + z \cdot \text{STD} \cdot (R+L)^{1/2} \).
- Expected level of stocks before an order arrives, \( z \cdot \text{STD} \cdot (R+L)^{1/2} \).
- The average inventory level is the average of two latter values and is determined by the following formula: \( (R \cdot \text{AVG} \cdot z) / 2 \cdot \text{STD} \cdot (R + L)^{1/2} \).

The explanation of symbols are as follow: The \( R \) is reorder point, \( L \) is lead time, \( \text{STD} \) is standard deviation, \( z \) is the service factor for safety stock, and \( \text{AVG} \) is average.

2.2. Support Vector Regression (SVR)

SVR is the regression method which developed from support vector machine (SVM) where it was created the hyperplane closest to an amount of data and becomes the objective. The small norm
hyperplane has collected distance from the data points to minimizing, shows in Figure 2. The $\varepsilon$-insensitive loss function can be pictured as a tube equals to calculate the accuracy where the environments of training data follows to the support vector learning.

![Figure 2. $\varepsilon$-sensitive loss function of SVR](image)

The regression estimated with SVR estimates a function according to a data set $(x_i, y_i)_n$ where $x_i, y_i$, and $n$ are input, output, and data points, respectively [12].

2.3. Adaptive Neuro-Fuzzy (ANFIS).

Takagi-Sugeno and Hayashi found the ANFIS method. It is formed by the parameters of antecedents to gradient descent, algorithm, and adjusts its following parameters by Least Square Estimator (LSE) algorithm. It calls the NN and the fuzzy system of learning procedure. [13]. If $X_i = A_i$ and $X_n = B_i$ then $f_i = p_i x_i + q_i x_n + r_i$ where $p_i, q_i$, and $r_i$ are parameters design to determine the training process [14]. Figure 3 shows the five layers, the first layers are adaptive. The linguistic layers show in (2) and (3) [15].

\[ O_{i,j} = \mu_{A_i}(x) \quad \text{for } i = 1, 2 \]  
\[ O_{i,j} = \mu_{B_i}(y) \quad \text{for } i = 3, 4 \]

Where $p_i, q_i, r_i$ are the parameters in the than-part (consequent part) of the first-order of Sugeno fuzzy model with five layers. Other layers follow the following formulas:

![Figure 3. ANFIS General Architecture](image)

Layer 2 consists of the nodes labeled H as follow (4):

\[ O_{2,i} = W_i = \mu_{A_i}(X)\mu_{B_i-2}(y) \quad i=1,2 \]  

In this part of N is fixed for every node (5).
\[ O_{3,i} = \overline{W_i} = \frac{w_i}{w_1 + w_2} \quad i=1,2 \quad (5) \]

Layer 4 shows the nodes from external the value of \( x \) and \( y \) (6).

\[ O_{4,i} = \overline{W_i} f_i (p_i x + q_i y + r_i) \quad (6) \]

Equation 6 shows \( \overline{W_i} \) which is the output of the layer 3 and \((p_i, q_i, r_i)\) are the parameters set.

Layer number 5 is the output of the sums to enter at coming signals (7).

\[ O_5^i = \sum W_i x f_i = \frac{\sum_i W_i x f_i}{\sum_i W_i} \quad (7) \]

2.4. Mean Square Error (MSE)

MSE is the calculation of the average square error with the formula in (8):

\[ MSE = \frac{\sum_{i=1}^{n}(y_i - F_i)^2}{n} \quad (8) \]

Where \( y_i \) is the actual value at time \( i \). \( F_i \) is the symbol of forecast value at the period then \( n \) is the period number.

3. Method

This study uses a quantitative method to process and analyze of the dataset. The variables are collected from price and demand dataset. They are generated by random normal before the variables entrance to the network. The output of methods is measured their SD and mean. This is useful to generate the variance of order in the beer game system. Moreover, we calculate the BWE based on the formula of beer game. Then, we generate the pattern of the BWE, then make its conclusion.

The data is chosen to be entered as input variables.

3.1. Support Vector Regression

Support vector regression used SteveGunn’s algorithm[16] then modified it. It calculated the svr and svr-output. The data was used 100 data, and it was separated into two parts. 90% was used for training and the rest of 10% was used for testing [17]. Then the data are poles between the border of insensitive (+ and -) close to the centerline to obtain the optimal solution.

The SteveGunn’s algorithm steps are as follow:

Step 1. \( \text{if (nargin < 3 | nargin > 6) \% Check the correct number of argument} \)

Step 2. \( \text{epsilon = svtol(UpbC) \% Tolerance for support vector detection} \)

Step 3. \( \text{Construct the kernel matrix} \)

\( \text{fprintf('Constructing ...\n');} \)
\( \text{H = zeros(n,n);} \)
\( \text{for i=1:n} \)
\( \text{for j=1:n} \)
\( \text{H(i,j) = svkernel(Metker,DataKu(i,:),DataKu(j,:));} \)
\( \text{end} \)
\( \text{end} \)

Step 4. Set up the parameter for the optimization problem

\( \text{switch lower(lossf)} \)
\( \text{case 'einsensitive',} \)
\( \text{Hb = [H -H; -H H];} \)
\( \text{c = [(insen*ones(n,1) - DataOk); (insen*ones(n,1) + DataOk)];} \)

Step 5. Set the bound: \( \boldsymbol{a} > 0 \) and \( \boldsymbol{a} \leq C \)

\( \text{vlb = zeros(2*n,1);} \)
\( \text{vub = UpbC*ones(2*n,1);} \)

Step 6. Set the starting point is 0

\( \text{x0 = zeros(2*n,1);} \)
Step 7. Set the constrain $Ax = b$
\[
\text{if neqcstr} \\
A = [\text{ones}(1,n) -\text{ones}(1,n)];, b = 0; \\
\text{else} \\
A = [], b = []; \\
\text{end}
\]
\[
\text{case 'quadratic'}, \\
Hb = H + \text{eye}(n)/(2*UpbC); \\
c = -DataOk; \\
vlb = -\text{le30}\text*ones(n,1); \\
vub = \text{le30}\text*ones(n,1);
\]
Step 8. Set the number of equality constrain (1 or 0)
\[
x0 = \text{zeros}(n,1);
\]
Step 9. Solve the optimization problem
\[
Hb = Hb+1e-10*\text{eye(size(Hb))}; \\
fprintf('Optimising ...\n'); \\
st = \text{cputime};
\]
Step 10. [alpha lambda how] = qp (Hb, c, A, b, vlb, vub, x0, neqcstr)
\[
fprintf('Execution time : %4.1f seconds\n',cputime - st);
\]
Step 11. Compute the number of Support Vectors
\[
svi = \text{find( abs(beta) > epsilon )}; \\
nsv = \text{length( svi )}; \\
fprintf('Support Vectors : %d (%3.1f\%)\n',nsv,100*nsv/n);
\]
Step 12. Implicit bias, $b_0$
\[
bias = 0;
\]
Step 13. Find the bias from an average of SVs with interpolant error $e$ with alphas: $0 < \alpha < C$.
\[
svii = \text{find( abs(beta) > epsilon & abs(beta) < (UpbC - epsilon))}; \\
\text{if length(svii) > 0} \\
bias = (1/\text{length(svii)})*\text{sum(DataOk(svii)} - \text{H(svii,svi)}*\text{beta(svi)}) \\
\text{else}
\]

3.2. ANFIS
The ANFIS follows the steps of the system in the software of Matlab, the system is drawn as follows:

1. Training Data
The training data, trnData, is a required argument to anfis, as well as to the ANFIS Editor GUI. Each row of trnData is a desired input/output pair of the target system you want to model. Each row starts with an input vector and is followed by an output value. Therefore, the number of rows of trnData is equal to the number of training data pairs, and, because there is only one output, the number of columns of trnData is equal to the number of inputs plus one.

2. Input FIS Structure
You can obtain the input FIS structure, fismat, from any of the fuzzy editors:

3. The FIS Editor
The Membership Function Editor
The Rule Editor from the ANFIS Editor GUI (which allows a FIS structure to be loaded from a file or the MATLAB workspace)
The command line function, genfis1 (for which you only need to give numbers and types of membership functions)
The FIS structure contains both the model structure, (which specifies such items as the number of rules in the FIS, the number of membership functions for each input, etc.), and the parameters, (which specify the shapes of membership functions).
There are two methods that anfis learning employs for updating membership function parameters: backpropagation for all parameters (a steepest descent method)

4. A hybrid method
This method is consisting of backpropagation for the parameters associated with the input membership functions, and least squares estimation for the parameters associated with the output membership functions.
5. The output

As an output, the training error decreases, at least locally, throughout the learning process. Therefore, the more the initial membership functions resemble the optimal ones, the easier it will be for the model parameter training to converge. Human expertise about the target system to be modeled may aid in setting up these initial membership function parameters in the FIS structure. The genfis1 function produces a FIS structure based on a fixed number of membership functions. This structure invokes the so-called curse of dimensionality, and causes excessive propagation of the number of rules when the number of inputs is moderately large, that is, more than four or five. Fuzzy Logic Toolbox software offers a method that provides for some dimension reduction in the fuzzy inference system: you can generate a FIS structure using the clustering algorithm discussed in Subtractive Clustering. To use the clustering algorithm, you must select the Sub. Clustering option in the Generate FIS portion of the ANFIS Editor GUI before the FIS is generated. This subtractive clustering method partitions the data into groups called clusters, and generates a FIS with the minimum number rules required to distinguish the fuzzy qualities associated with each of the clusters.

4. Results and discussion

From this study, we prepared the data to check the missing data and synchronized the unit. The variables are used in this study, Price ($D_1$) and Demand ($D_2$) of tin milk. Then the data put between -1 and 1 from normalized.

The fluctuations of data tend to form as time series dataset from 1 to 100 data. Then, we used SVR and ANFIS to measure the BWE.

4.1. SVR for Tin Milk Industry

The SteveGunn’s algorithm code and it is modified to calculate the output of SVR function in the program code. Ninety percent apply for training, then the rest for testing. The $\varepsilon$ parameter manages the large of the $\varepsilon$-insensitive area, where it is used to fit the training dataset. The support vector (SV) machine solves to nonlinear function in the optimization problem. We use the value of $\varepsilon$ with 1 (because the value of $\varepsilon \geq 0$), and it can be affected by the amount of SV to build the regression function [18].

4.2. ANFIS for Tin Milk industry

The data run with ANFIS tool to get the training output of the fuzzy inference system. It means the square error is recorded by the training error in every step of the epoch. The statistical indicator is used for measuring the model performance. The results from these methods are illustrated in Table 1.

| Method | Statistics  | MSE         |
|--------|-------------|-------------|
|        | SD          | Mean        |
| ANFIS  | 0.2075      | 0.5556      | $5.178 \times 10^{-4}$ |
| SVR    | 0.2201      | 0.6810      | $2.255 \times 10^{-2}$ |

Table 1 shows that the performance of ANFIS is a better model compared to the SVR, it indicates the MSE value. SD and variance tend to decrease from SVR and ANFIS. The statistical data are used to get the variance of order in the supply chain where it is obtained from the beer game system. The Bullwhip Effect is the amplification of variance order and demand. The variance order is obtained from the beer game software to get the result of statistical output from each member of the supply chain as shown in the Table 2. Moreover, the ANFIS is found to be a better performance compared to the SVR. It can be shown the results that from a retailer, wholesaler, distributor, and factory tend to increase.
Table 2. The bullwhip effect of three methods

| Method | BWE Retailer | BWE Wholesaler | BWE Distributor | BWE Factory | Total BWE |
|--------|--------------|----------------|----------------|-------------|-----------|
| ANFIS  | 0.354        | 40.909         | 72.673         | 95.047      | 208.773   |
| SVR    | 2.103        | 232.86         | 421.7          | 580.81      | 1237.263  |

5. Conclusion

This study explores the tin milk data in Indonesia. We argue, to produces a better performance of forecasting method should lead to minimize the BWE in the supply chain. The oscillations of demand can contribute the amplification performance in the supply chain. It may be the causes of uncontrollable inventory, a quality control problem, and lack of customer service. This risk can increase cost expenditure for supply chain member. This study try, the exploration the methods of ANFIS and SVR to measure the bullwhip effect in across of the supply chain. The result showed that the ANFIS is better than SVR where it contributes the smallest BWE. The total results of the BWE in each member for the ANFIS method are 0.354, 40.909, 72.673, and 95.047 from a retailer, wholesaler, distributor, and the factory respectively and total BWE is 208.773. In this study, there are two significant conclusions. First, better forecasting performance provides the smaller amplification of BWE in the member of a supply chain. The second conclusion is the ANFIS is better than SVR.

References

[1] H. L. Lee, V. Padmanabhan, and S. Whang, "Information distortion in a supply chain: the bullwhip effect," Management science, vol. 50, pp. 1875-1886, 2004.
[2] W.-C. Wang, K.-W. Chau, C.-T. Cheng, and L. Qiu, "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series," Journal of Hydrology, vol. 374, pp. 294-306, 2009.
[3] H. Drucker, D. Wu, and V. N. Vapnik, "Support vector machines for spam categorization," IEEE Transactions on Neural Networks, vol. 10, pp. 1048-1054, 1999.
[4] M. Sugeno and G. Kang, "Structure identification of fuzzy model," Fuzzy Sets and Systems, vol. 28, pp. 15-33, 1988.
[5] M. Themistocleous, Z. Irani, and P. E. Love, "Evaluating the integration of supply chain information systems: A case study," European Journal of Operational Research, vol. 159, pp. 393-405, 2004.
[6] H. L. Lee, V. Padmanabhan, and S. Whang, "Information distortion in a supply chain: The bullwhip effect," Management science, vol. 43, pp. 546-558, 1997.
[7] Z. Isa, A. R. Alsayed, and S. S. Kun, "Review Paper on economic growth-aggregate energy consumption nexus," International Journal of Energy Economics and Policy, vol. 5, 2015.
[8] E. Fradinata, S. Suthummanon, and W. Suntiamornrut, "Reducing the bullwhip effect from signal demand of hybrid artificial neural network models of supply chain in Indonesia," International journal of advanced and applied sciences, vol. 4, pp. 64-75, 2017.
[9] H. L. Lee, V. Padmanabhan, and S. Whang, "The bullwhip effect in supply chains," MIT Sloan Management Review, vol. 38, pp. 93, 1997.
[10] R. N. Boute and M. R. Lambrecht, "Exploring the bullwhip effect by means of spreadsheet simulation," INFORMS Transactions on Education, vol. 10, pp. 1-9, 2009.
[11] D. Simchi-Levi, E. Simchi-Levi, and P. Kaminsky, Designing and managing the supply chain: Concepts, strategies, and cases: McGraw-Hill New York, 1999.
[12] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," Statistics and computing, vol. 14, pp. 199-222, 2004.
[13] D. Nauck and R. Kruse, "A neuro-fuzzy method to learn fuzzy classification rules from data," *Fuzzy Sets and Systems*, vol. 89, pp. 277-288, 1997.

[14] A. Hasiloglu, M. Yilmaz, O. Comakli, and İ. Ekmekci, "Adaptive neuro-fuzzy modeling of transient heat transfer in circular duct air flow," *International journal of thermal sciences*, vol. 43, pp. 1075-1090, 2004.

[15] A. K. Palit and D. Popovic, "Computational intelligence in time series forecasting," ed, 2006.

[16] S. R. Gunn, "Support vector machines for classification and regression," *ISIS technical report*, vol. 14, pp. 5-16, 1998.

[17] E. Fradinata, S. Suthummanon, and W. Suntiamorntut, "Initial Optimal Parameters of Artificial Neural Network and Support Vector Regression," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, 2018.

[18] B. Santosa, "Data Mining; ," in *Data mining techniques for business application* vol. first edition, ed: Yogyakarta: Graha Ilmu, 2007, pp. 45-50.