

PAPER • OPEN ACCESS

Forecasting tourist visits using data decomposition technique and learning optimization of artificial neural network

To cite this article: S Herawati and M Latif 2021 *J. Phys.: Conf. Ser.* **1811** 012092

View the [article online](#) for updates and enhancements.

You may also like

- [Exploring the impacts of tourism and weather on water consumption at different spatiotemporal scales: evidence from a coastal area on the Adriatic Sea \(northern Italy\)](#)
Filippo Mazzoni, Valentina Marsili, Stefano Alvisi et al.
- [Detection of coronary artery disease using multi-modal feature fusion and hybrid feature selection](#)
Huan Zhang, Xinpei Wang, Changchun Liu et al.
- [Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features](#)
Faiq Ahmad Khan, Anam Abid and Muhammad Salman Khan

PRIME
PACIFIC RIM MEETING
ON ELECTROCHEMICAL
AND SOLID STATE SCIENCE

HONOLULU, HI
Oct 6–11, 2024

Abstract submission deadline:
April 12, 2024

Learn more and submit!

Joint Meeting of
The Electrochemical Society
•
The Electrochemical Society of Japan
•
Korea Electrochemical Society

Forecasting tourist visits using data decomposition technique and learning optimization of artificial neural network

S Herawati^{1*} and M Latif²

¹Department of Information Systems, University of Trunojoyo Madura

²Department of Mechatronics Engineering, University of Trunojoyo Madura

*Corresponding author : sriherawati@trunojoyo.ac.id

Abstract. Estimates of tourist visits is very important to determining policy and decision making. This study proposed a new method for forecasting tourist visits. A case study was conducted at a tourist spot in Sumenep, Indonesia. The model proposed is data decomposition and optimization of learning against tourist visits data. Data decomposed use the Ensemble Empirical Mode Decomposition (EEMD) method, then data learning use the Feedforward Neural Network (FNN) which was optimized using the Polak-Ribiere Conjugate Gradient (PCG). The two methods are integrated to produce accurate forecasts. Several patterns of learning data were carried out in this experiment. The results of this method show good performance results as measured used RMSE and MSE. **Keyword** : EEMD, PCG, Forecasting, Tourist Visits

1. Introduction

Tourism is used as an important sector to support the economic growth of a government. Tourism planning and management is carried out to attract tourists, both domestic and foreign tourists. Lack of preparation of food supplies, hotels and other tourist needs makes them disappointed and evaluated poorly. On the other hand, if the excess will suffer a big loss for an industry [1]. The government needs an estimate to formulate development and infrastructure policies that support the tourism industry. Forecasting tourist visits is important to help determine policies and decision-making by the government and the tourism industry. Therefore, one of the critical issues in the tourism industry is the development of an accurate tourism demand forecasting model [2].

Forecasting tourist visits is done by processing and analyzing the time series data patterns. These data are non-stationary, nonlinear and difficult to use to develop accurate forecasting models, because tourist visits are influenced by many factors, including the holiday season [3], economy [4], weather [5], plague [6], facilities [7].] and other factors [8]. Many researchers use time series models to build forecasting systems. There are various approaches to forecasting tourist visits by considering past time series data patterns [9]. Several studies have investigated the data decomposition technique approach to forecasting tourist visits [9] [10]. The ability of decomposition techniques has been proven in various fields, for example crude oil forecasting [11], network traffic [12], wind speed [13] and electrical load [14].

A popular decomposition technique proposed by Huang et al. [15] namely the Empirical Mode Decomposition (EMD) which was later refined into the Ensemble Empirical Mode Decomposition (EEMD) by Wu and Huang [16-17]. This method has been shown to be flexible enough to extract



signals that have nonlinear and non-stationary characteristics. EEMD will decompose the data into several IMFs and a residue. This method is at the core of this research.

This study has proposed a novel model for forecasting tourist visits using a combination of EEMD and artificial neural network (ANN). The Learning ANN used is a type of feed-forward neural network (FNN) that has been optimized using the Polak-Ribiere Conjugate Gradient (PCG). The PCG algorithm has advantages in computation speed and high quality compared to other learning algorithms [17]. We use the case of tourist visits in Sumenep district, Indonesia to verify proposed model performance .

2. Methods

This study has proposed a new model to produce accurate forecasting of tourist visits using data decomposition techniques and artificial Neural Networks that have been optimized using PCG. The forecasting model is shown in Figure 1.

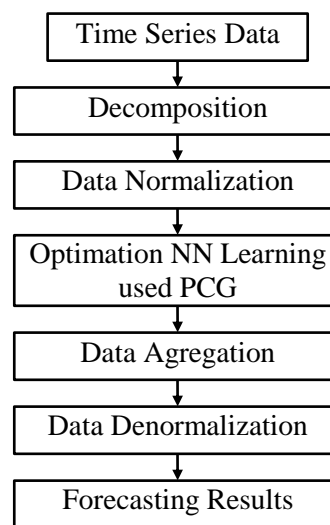


Figure 1. Framework of the proposed forecasting model.

2.1. Decomposition Technique

The data decomposition technique uses the EEMD method. That is a modification of the EMD. The EEMD is used to correct the weakness of EMD which results in a broken signal so that it displays mixed mode. The data decomposition technique using EEMD consists of the following steps [16];

Step 1: Initialize the ensemble number (M)

Step 2: Generate white noise using (1);

$$n_{m(t)} = Nstd * randn(x_{(t)}) \tag{1}$$

$n_{m(t)}$ is the m -th white noise, $Nstd$ is the standard deviation of the noise, and $randn(x_{(t)})$ is normally distributed pseudo-random numbers based on tourist visit data $x_{(t)}$.

Step 3: Testing the addition of the white noise signal (2);

$$x_{m(t)} = x_{(t)} + n_{m(t)} \tag{2}$$

$x_{(t)}$ is touris visits time series, $x_{m(t)}$ is the test result of the white noise addition.

Step 4: Identify all local maximum and minimum values at $x_{m(t)}$ and use the cubic spline function to generate the top and bottom covers.

Step 5: Find the mean $m_{1(t)}$ for the top and bottom covers. Then, use (3) to calculate the error $e_{1(t)}$ between the signal and the mean;

$$e_{1(t)} = x_{m(t)} - m_{1(t)} \tag{3}$$

Step 6: If $e_{1(t)}$ meets IMF criteria, then $c_{1(t)} = e_{1(t)}$ is the first IMF component of the signal. If not, change $x_{m(t)} = e_{1(t)}$ and do step 4. The IMF has two criteria, that is, (1) for the entire data series, the maximum difference between the number of extreme values (the sum of local maximum and minimum values) and the number of zero intersections is one. (2) the mean value of the covers must be equal to zero at each point.

Step 7: For the residual data $r_{1(t)}$ expressed as (4);

$$r_{1(t)} = x_{m(t)} - c_{1(t)} \tag{4}$$

assuming the residue as a new signal, then repeat steps 4 through step 7 for every n to $c_{n(t)}$ or $r_{n(t)}$ less than the predetermined value or the residue being a monotonous function of the unextracted IMF.

Step 8: After the decomposition process, data can be represented as the sum of all IMF and residues as shown in (5);

$$x_{m(t)} = \sum_{i=1}^n c_{i(t)} + r_{n(t)} \tag{5}$$

n is the number of IMFs, $r_{n(t)}$ is the final residual signal, and $c_{i(t)}$ is the i -th IMFs.

Step 9: If $m < M$, do $m = m + 1$ and repeat step 3 through step 7 used different white noise. Process will stop when $m = M$.

Step 10: Find the ensemble mean for each IMFs and the residue in experiment M using (6) dan (7);

$$\bar{c}_i = \frac{1}{M} \sum_m^M c_{i,m} \quad \rightarrow \quad i = 1, \dots, n \tag{6}$$

$$\bar{r}_n = \frac{1}{M} \sum_m^M r_{n,m} \quad \bar{r}_n = \frac{1}{M} \sum_m^M r_{n,m} \tag{7}$$

2.2. Data Normalization

Data normalization is carried out as an activation function of the input values in neural network training. Data normalization is expressed using (8);

$$x' = y_{\min} + (y_{\max} - y_{\min}) * \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) \tag{8}$$

x' is normalized data, y_{\min} and y_{\max} are the minimum and maximum values of the activation function value range, x is the actual time series data to be normalized, x_{\min} dan x_{\max} are the minimum and maximum values of the actual time series data respectively.

2.3. Learning Neural Network with PCG optimization

Data learning was carried out using FNN which was optimized with the Polak-Ribière Conjugate Gradient. PCG is one of the neural network learning algorithms proposed by Polak and Ribière. The search for the negative value of the grid gradient starts from the first iteration and is based on the direction of the conjugation. The FNN learning algorithm with PCG optimization is as follows;

Step 1: Initialize all weights using small random numbers

Step 2: do step 2 through 13 if the termination conditions are not met.

Step 3: If the input layer receives data, then pass it on to the hidden layer.

Step 4: Find all hidden layer output h_j ($j = 1, 2, \dots, p$);

$$h_{-net j} = v_{j0} + \sum_{i=1}^n x_i v_{ji} \tag{9}$$

$$h_j = f(h_{-net j}) = \frac{1}{1 + e^{-h_{-net j}}} \tag{10}$$

Step 5: Find all the outputs on the output layer y_k ($k=1, 2, \dots, m$);

$$y_net_k = w_{k0} + \sum_{j=1}^p h_j w_{kj} \tag{11}$$

$$y_k = f(y_net_k) = \frac{1}{1 + e^{-y_net_k}} \tag{12}$$

Step 6: Find the error factor in the output layer based on the difference between the actual error value and the forecast value;;

$$e_k = (t_k - y_k) f'(y_net_k) \tag{13}$$

$$e_k = (t_k - y_k) y_k (1 - y_k) \tag{14}$$

Step 7: Find the error factor in the hidden layer based on the error factor of the layer above;

$$e_net_j = \sum_{k=1}^m e_k - w_{kj} \tag{15}$$

$$e_j = e_net_j f'(h_net_j) \tag{16}$$

$$e_j = e_net_j h_j (1 - h_j) \tag{17}$$

Step 8: Find the gradient on the output layer of the defined objective function;

$$m_{k+1} = \frac{1}{N} \sum_{n=1}^p e_{nk} y_{nk} \tag{18}$$

Step 9: Find the gradient on the hidden layer;

$$m_{j+1} = \frac{1}{N} \sum_{n=1}^p e_{nj} y_{nj} \tag{19}$$

Step 10: Find the parameter β for all neurons in hidden layer and output layer;

$$\beta_{t+1} = \frac{m_{t+1}^T (m_{t+1} - m_t)}{m_t^T m_t} \tag{20}$$

Step 11: Find the directions for all neurons in the hidden layer and the output layer;

$$d_{t+1} = -m_{t+1} + \beta_t d_t \tag{21}$$

For the first direction:

$$d_t = -m_t \quad d_t = -g_t \tag{22}$$

Step 12: Find the α parameter for all neurons in the hidden layer and the output layer;

Step 13: Update the weight using (23);

$$w_{t+1} = w_t + \alpha_{t+1} d_{t+1} \tag{23}$$

2.4. Data Aggregation

Data aggregation serves to combine data that has been decomposed. The decomposition data consists of several IMFs and a residue. Data aggregation was carried out using the adaptive linear neural network (Adaline) method. Adaline method is written in (24);

$$y = \sum_{i=1}^n x_i \cdot w_i + b \tag{24}$$

2.5. Data Denormalization

Denormalization is used to restructure Adaline's output into real data values. Denormalization is calculated using (8) to find the x .

Two methods are used to evaluate the performance of the proposed forecasting model, namely Mean Square Error (MSE) and Root Mean Squared Error (RMSE). The MSE and RMSE equations are shown in (25) and (26).

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - x_t)^2 \tag{25}$$

$$RMSE = \sqrt{\sum_{t=1}^n (y_t - x_t)^2} \quad (26)$$

n is number of actual data, x_t is actual data and y_t forecast data.

3. Results And Discussions

3.1. Data Collection

As an illustration and verification of the forecasting model, an empirical study was conducted using time series data for tourist visits in Sumenep Regency, Indonesia. The sample data used is from January 2015 to December 2019 as shown in Figure 2. The data used for training is 70% and the rest is used as testing.

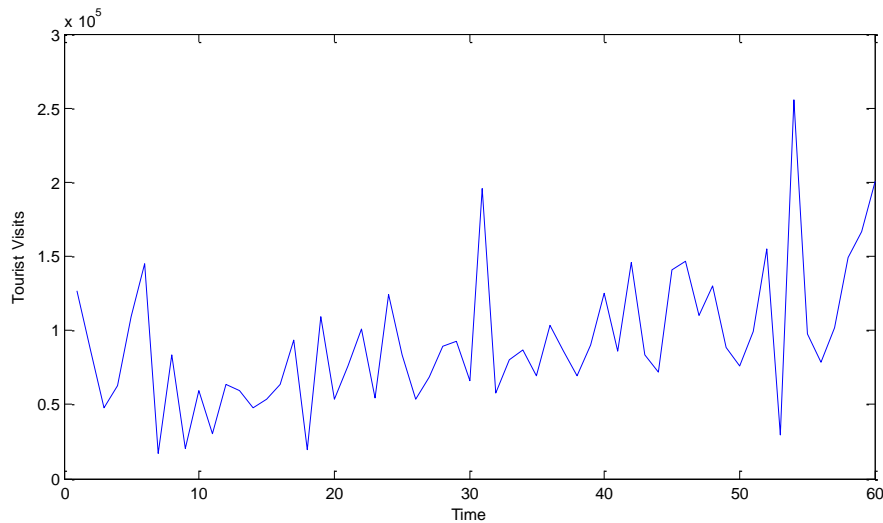


Figure 2. Monthly tourist visits from 2015 - 2019.

3.2. Forecasting Result

This experiment uses the number of ensembles of 100 with a standard deviation of 0.2 referring to research conducted by [16]. Experimental data were decomposed using EEMD. The EEMD process produces four IMFs and one Residue as shown in Figure 3. All IMFs are arranged from highest to lowest frequency to represent periodic changes from high time variance or disturbance to long periods. The value of each IMFs and residue is then normalized to be converted into a range of values between 0 and 1. Normalization aims to fulfill the requirements of the binary sigmoid activation function in the experiment. The results of normalization are shown in Figure 4, it can be seen that the shape of the signal frequency is the same as Figure 3, but the value is different.

Training and testing using FNN with PCG learning uses the following parameters; The number of iterations is 10000 times, the learning rate is 0.1, and the error tolerance is 0.0001. The experiment was carried out in two scenarios, first, changing the number of hidden layer neurons in the FNN and second, changing the number of input neurons. The scenario by changing the number of hidden layers is done with a data pattern of multiples of 5 then the forecast performance is measured using RMSE and MSE. The data pattern and performance of this scenario are shown in table 1. The experimental results show that the 3-30-1 data pattern has the best performance with RMSE is 0.15646 and MSE is 0.02448. After the first scenario was carried out, the hidden layer neuron data pattern that had the best performance value was used in the second scenario. In the second scenario, the FNN network architecture is determined by changing the input layer pattern. The data patterns were applied to the IMF and residues, respectively. The data pattern and performance of this scenario are shown in table 2. The experimental results show that the 4-30-1 data pattern is the best performance with RMSE is 0.14592 and MSE is 0.02129. After the learning process, the model validation is carried out on the

tourist visit data. The comparison between actual data and tourist visit forecasting data is shown in Figure 5.

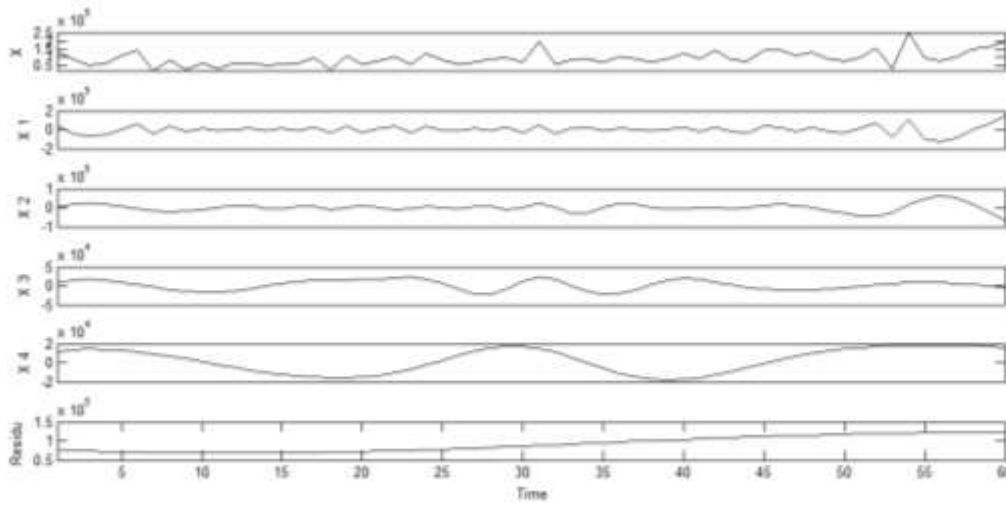


Figure 3. Result of ensemble empirical mode decomposition

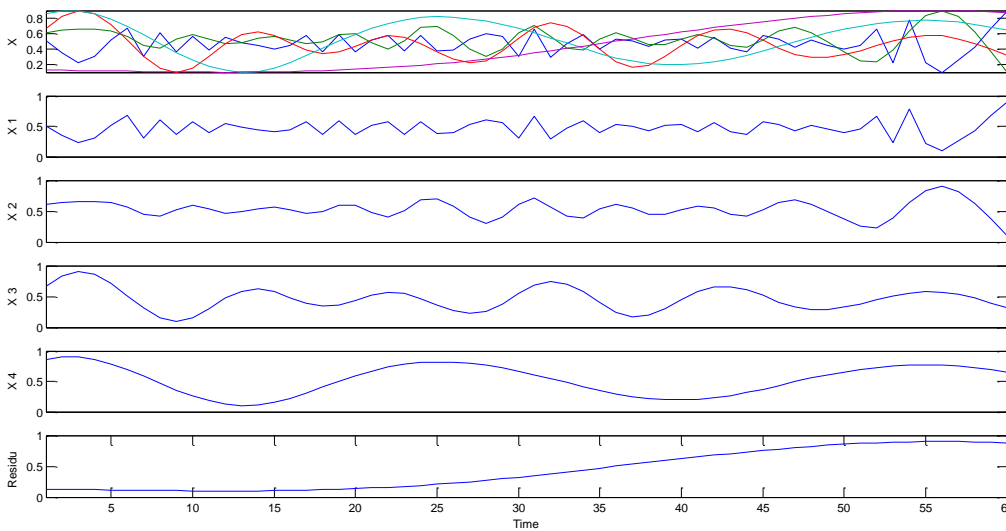


Figure 4. Result of normalization

Table 1. Performance by changing hidden layer

Pola data	RMSE	MSE
3-5-1	0.17254	0.02977
3-10-1	0.16355	0.02675
3-15-1	0.17356	0.03012
3-20-1	0.16031	0.02570
3-25-1	0.15739	0.02477
3-30-1	0.15646	0.02448

Table 2. Performance by changing hidden layer

Pola data	RMSE	MSE
3-30-1	0.15646	0.02448
4-30-1	0.14592	0.02129
5-30-1	0.16969	0.02880
6-30-1	0.18248	0.03330

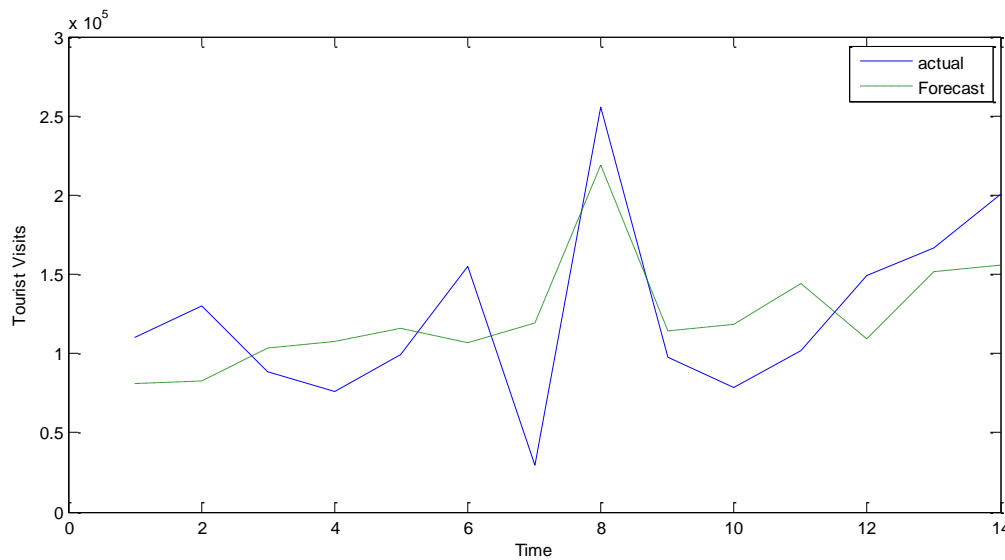


Figure 5. Forecasts of Tourist Visits to Sumenep

4. Conclusions

This study has proposed a new model to produce accurate forecasting of tourist arrivals using data decomposition techniques and artificial Neural Networks that have been optimized using PCG. The model was tested on tourist visits data in Sumenep Regency, Indonesia. EEMD produces 4 IMFs and one residue. FNN with PCG learning is carried out to find the best pattern by changing the hidden layer neurons and the input layer. The best performance produced is a 4-30-1 data pattern (4 input neurons, 30 hidden neurons, and 1 output) with RMSE is 0.14592 and MSE is 0.02129.

5. References

- [1] Palmer A, José Montaña J and Sesé A 2006 Designing an artificial neural network for forecasting tourism time series *Tour. Manag.* **27** 781–90
- [2] Xie G, Qian Y and Wang S 2020 A decomposition-ensemble approach for tourism forecasting *Ann. Tour. Res.* **81** 102891
- [3] Tol R S J, Bigano A and Hamilton J M 2011 The Impact of Climate on Holiday Destination Choice *SSRN Electron. J.*
- [4] Holik A 2016 Relationship of Economic Growth with Tourism Sector *JEJAK J. Ekon. dan Kebijak.* **9** 16–32
- [5] Day J, Chin N, Sydnor S and Cherkauer K 2013 Weather, climate, and tourism performance: A quantitative analysis *Tour. Manag. Perspect.* **5** 51–6
- [6] Haryanto T 2020 Editorial: Covid-19 Pandemic and International Tourism Demand *J. Dev. Econ.* **5** 1
- [7] Baker D M and Unni R 2019 Characteristics And Intentions Of Cruise Passengers To Return To The Caribbean For Land-Based Vacations *J. Tour. -studies Res. Tour.*
- [8] Martins L F, Gan Y and Ferreira-Lopes A 2017 An empirical analysis of the influence of macroeconomic determinants on World tourism demand *Tour. Manag.* **61** 248–60

- [9] Li C, Ge P, Liu Z and Zheng W 2020 Forecasting tourist arrivals using denoising and potential factors *Ann. Tour. Res.* **83** 102943
- [10] Xie G, Qian Y and Wang S 2020 A decomposition-ensemble approach for tourism forecasting *Ann. Tour. Res.* **81** 102891
- [11] Herawati S and Latif M 2018 Crude Oil Price Forecasting Using Ensemble Empirical Mode Decomposition and Generalized Regression Neural Networks *Adv. Sci. Lett.* **23** 12413–6
- [12] Xie G, Qian Y and Wang S 2020 A decomposition-ensemble approach for tourism forecasting *Ann. Tour. Res.* **81** 102891
- [13] Guo Z, Zhao W, Lu H and Wang J 2012 Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model *Renew. Energy* **37** 241–9
- [14] Yaslan Y and Bican B 2017 Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting *Meas. J. Int. Meas. Confed.* **103** 52–61
- [15] Huang N E, Shen Z, Long S R, Wu M C, Snin H H, Zheng Q, Yen N C, Tung C C and Liu H H 1998 The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* vol 454 pp 903–95
- [16] Wu Z and Huang N E 2009 Ensemble empirical mode decomposition: A noise-assisted data analysis method *Adv. Adapt. Data Anal.* **1** 1–41
- [17] Simanullang NH, Rajagukguk J. Learning Management System (LMS) Based On Moodle To Improve Students Learning Activity. *JPhCS.* 2020 Feb;1462(1):012067.
- [18] Ghasemi-Fard M, Ansari-Asl K, Albera L, Kachenoura A and Senhadji L 2013 A low cost and reliable Polak-Ribiere conjugate gradient deflation ICA algorithm for real signals *2013 21st Iranian Conference on Electrical Engineering, ICEE 2013*

Acknowledgments

The authors wish to thank University of Trunojoyo Madura for funding this research.