Malaria Incidence Forecasting from Incidence Record and Weather Pattern Using Polynomial Neural Network

Anditya Arifianto, Ari Moesriami Barmawi, and Agung Toto Wibowo

Abstract—**Malaria affects over 100 million persons worldwide with approximately 2,414 deaths a day in average each year. Indonesia is on the third highest position in the number of malaria incident in South East Asia, with 229,819 confirmed cases and 432 deaths only at 2010. Previous work has demonstrated the potential of neural networks in predicting the behavior of complex, non-linear systems. GMDH Polynomial Neural Network was applied in a great variety of areas for data mining and knowledge discovery, forecasting, systems modeling, optimization, and pattern recognition. Study has also shown the close relation between Malaria incidence and weather pattern. This paper proposed a modified GMDH Polynomial Neural Network to reduce the learning time and computation while maintaining the accuracy in predicting Malaria incidence by relating it to weather pattern. Based on the experiments, it was proven that the modified GMDH PNN was able to reduce the learning time by 72% and improve the accuracy into 88.02% compared to the original GMDH PNN.**

Index Terms—**Malaria, prediction, weather pattern, polynomial neural network.**

I. INTRODUCTION

Malaria affects over 100 million persons worldwide each year. The impact of malaria can cause about 2,414 deaths a day [1]. In 2000-2010, around South East Asia (SEA) Region, it was reported that malaria incidence remained around $2.30 - 3.08$ million cases and deaths around $2,423 - 6,978$ annually. Indonesia located in SEA has a great number of incidents, the third highest position of malaria incident, with number of confirmed cases of 229,819 cases reported and 432 deaths [2].

Studies of malaria epidemics in these areas have shown their association with weather condition to the *anopheles* mosquito as the medium [1], [3]. The *Anopheles* mosquitoes like to live in areas that have an abundance of natural water because they need this natural water to breed. Some breed in shady areas such as forests, while others breed in open fields where there is plenty of sunlight. This is seen in the direct correlation between an abundance of *Anopheles* mosquitoes and rainfall, temperature, and vegetation density [1], [3].

Banggai regency, Central Sulawesi, is one of the malaria prone areas because of its agricultural areas, rice fields and marshes. 16 out of 19 sub-districts are at high risk with number of malaria transmission above 5 cases per 1000 population per year. Due to the severe health impact of malaria epidemics there is a growing need for methods that allow forecasting, early warning and timely case detection in areas of unstable transmission, so that more effective control measures can be implemented

Malaria incidence forecasting is a problem of time series prediction, and a Time series prediction involves the determination of an appropriate model, which can encapsulate the dynamics of the system, described by the sample data. Previous work has demonstrated the potential of neural networks in predicting the behavior of complex, non-linear systems. In particular, the class of polynomial neural networks has been shown to possess universal approximation properties, while ensuring robustness to noise and missing data, good generalization and rapid learning [4].

Group Method of Data Handling (GMDH) Polynomial Neural Network has been applied in a great variety of areas for data mining and knowledge discovery, forecasting and systems modeling, optimization and pattern recognition. Inductive GMDH algorithms give possibility to find automatically interrelations in data, to select optimal structure of model or network and to increase the accuracy of existing algorithms. And Because of that, for this research, to make a prediction system of malaria incidence, a self-organizing neural network known as GMDH Polynomial Neural Network will be used [5].

Based on the rationale above, it can be identified that Malaria is a serious disease in Banggai regency. A system to predict the Malaria outbreak is needed so that the Ministry of Indonesia can make a precaution for the impact. However, the Health Department of Ministry of Indonesia has not got a forecasting method with a good performance.

The existing forecasting method owned by the ministry is limited to study of auto regression such as ARIMA. Most standard ARIMA based programs as a linear model for forecasting has a significant weakness to outlier detection. That is, when an outlier occurs not alone do the parameters have to be re-estimated but it may also be the case that order of the ARIMA model has also changed at that point. With the urgency of the problem which has a time series properties, a more adaptive method is needed as the auto regression technique has a limited adaptation toward a further prediction.

Many researches about time series prediction only use one property as input and output. However, to correlate the malaria incidence and weather factors, the system is faced with multiple data series as input. Although GMDH PNN was considered has a fast learning time with its self-organizing properties, the standard GMDH PNN still perform excessive calculation and complexity by blindly try all possible architecture in self-organizing training process.

Thus, a modification to reduce the complexity and to guide the learning process is needed so the learning will proceed faster while maintaining its accuracy.

The objective of this work was to study the relation among the weather factors and malaria incidence to develop a predictive model to forecast the Malaria incidence and to make a modification to GMDH Polynomial Neural Network that will reduce the complexity and training time while maintains the prediction performance compared to the original GMDH PNN.

II. RELATED WORK

Many studies about Malaria forecasting based on monthly case reports and environmental factors have been conducted using autoregressive method and artificial neural network [1], [3], [6].

The result shows that ARIMA manages to produce a prediction model that matches with behavior of malaria incidence. However, the process of prediction model building itself is a matter of trial and error. Furthermore, the autoregressive method shows that the model might not have had enough statistical power to yield significant values or be long enough in duration to capture seasonality [6].

Previous work also demonstrates the potential of neural networks in predicting the behavior of complex, non-linear systems [4]. However, a simple multilayer perceptrons (MLPs), such as the back propagation learning algorithm, has some weakness such as slow learning properties which utilize computationally intensive training algorithms and can get trapped in local minima [7].

Group Method of Data Handling (GMDH) Polynomial Neural Network is applied in a great variety of areas for data mining and knowledge discovery, forecasting and systems modeling, optimization and pattern recognition [5]. Inductive GMDH algorithms give possibility to find automatically interrelations in data, to select optimal structure of model or network and to increase the accuracy of existing algorithms.

III. ANOPHELES MOSQUITO BREEDING

The *Anopheles* mosquito is known as the medium transmission of malaria. Studies have shown that mosquito's life cycle is affected by many weather factors such as temperature, rainfall, and precipitation day[3]. The productivity of the mosquito varies depending on the weather. The *Anopheles* mosquitoes lay their eggs at temperature about 20°C-30°C. The eggs hatch in the water in 1 to 3 days at the temperature of 30°C, but at the temperature of 16°C the eggs hatch in 7 days. The average of optimum temperature for *Anopheles* is 25-27°C, while their growth stops in temperature less than 10°C or above 40°C. And their speed in regrowth and biological event, such as length of pre-adult, speed of digesting the blood, and biting frequency, depend on their metabolism speed which is affected by temperature.

The rains and the humidity increase the mosquitoes' natural inhabitant to lay their eggs. More rain means more water for mosquitoes to lay their eggs. The availability of water for a long time will keep the eggs until they hatch and become adult mosquitoes. The humidity is affecting the mosquitoes' longevity. Mosquitoes are breathing using air pipes which is called trachea, using the holes in their bodies called spiracle. With spiracle that has no control mechanism, in low humidity, the evaporation inside the mosquitoes will arise. In low humidity (less than 60%) their lifespan will reduce. Length of daylight affects the productivity of the mosquitoes. Mosquitoes only fly in low ray intensity in less than 20 Ft-cd, and mosquitoes' larvae can survive better in dark container

IV. GMDH POLYNOMIAL NEURAL NETWORK

High-level system modeling technique has been much attention recently, and all the details of the algorithms and methodologies are faced with nonlinear systems, high-dimensional problems, accuracy improvement, and generalization ability for the advanced techniques. In this case, nonlinear models can overcome some of these problems. But the weakness of nonlinear models are nonlinear models require very large data. When the complexity of the system to be modeled increased, both the experimental data and the domain knowledge needed to complete an efficient design procedure.

Polynomial Neural Network (PNN) is a GMDH type algorithm which is categorized as "self-organizing" networks. Each neuron in PNN called Partial Description (PD). The output of each PD is obtained using some type of order polynomial computations such as linear equations, quadratic equations, and even cubic equations of the input variables [8].

PNN can be viewed as a generalization of linear classification using the input not only of the measurement features of the input patterns, but also a polynomial of measurements. Ivakhnenko makes neurons become more complex with the unit displays the polynomial transfer function. Interconnection between layers of neurons simplified and automated algorithms for structural design and weight adjustment neurons developed.

Neuron-GMDH PNN has two inputs and output a polynomial combination of the two inputs so GMDH-PNN network forms a polynomial combination of input components. The general formula is biquadratic function as shown in Eq. (1) and Fig. 1

$$y = p_0 + p_1 x_1 + p_2 x_2 + p_3 x_1 x_2 + p_4 x_1^2 + p_5 x_2^2 \tag{1}$$
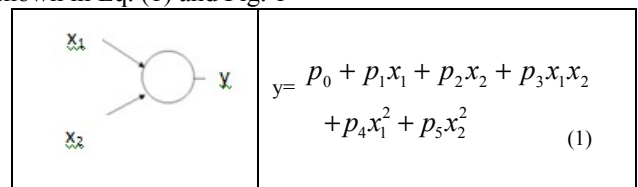
Fig. 1. Neuron of GMDH PNN

The first layer (far left) in Fig. 2 represents the input variables. Each neuron in the second layer takes the input from two input variables. Neurons in the third layer takes the input from two neurons in the previous layer, it went through each layer. The final layer (far right) draws two inputs from the previous layer and produces a single value which is the output of the network [9].

A. Data Preprocessing

The data processing in this study was a set of process applied to past malaria incidence and weather data before entering the main process of the system. It was aimed to get a proportional form and condition for the next processes. Data Preprocessing consists of Normalization and generate data series process. The weather factors consist of rainfall, precipitation day, minimum temperature, maximum temperature, average temperature, average humidity, and average length of daylight from 2004-2009.
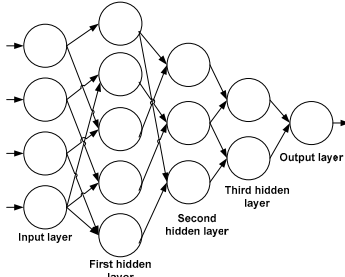

Fig. 2. GMDH PNN Architecture

The data used is still in the range of different values. Therefore, the normalization process is conducted to rescale all the data to the numerical range of values [0, 1]. Equation 2 is the equation used in the normalization of data [**10**].

$$x' = \left(\left(R_a \quad R_b \right) \frac{x - \min 2 Norm}{\max 2 Norm \; - \; \min 2 Norm} \right) + R_b \mathfrak{A} \quad (2)$$

where:

x' = normalized data

x = raw data

R_a = upper range of normalization range

R_b = lower range of normalization range

min2Norm = minimum raw value - margin

max2Norm = maximum raw value + margin

Considering not only the known values, but also the future values (those to be forecasted), the maximum and minimum limits for normalizing (max2norm, min2norm respectively) cannot be just the maximum (max) and minimum (min) known time series values. A margin from max and min has to be set if future values are higher or lower than they already are [11]. Eq. (3) and Eq. (4) calculate the addition of margin value given to the minimum and maximum value of current data to address the possibility of a value greater or smaller value in the future

$$\min 2 Norm = \min Data - (0.1 \times (\max Data - \min Data)) \quad (3)$$

$$\max 2 Norm = \max Data + (0.1 \times (\max Data - \min Data)) \quad (4)$$

where:

Min Data = minimum raw value

Max Data = maximum raw value

The provided data was available from 2004-2009. For the training process, the 6 year data was divided into three categories: Training Data, Validation Data, and Testing Data with ratio 4:1:1 year of data.

Like all mosquitoes, anophelines go through four stages in their life cycles: egg, larva, pupa, and imago. The first three stages are aquatic and last 5–14 days, depending on the species and the ambient temperature. The adult stage is when the female *Anopheles* mosquito acts as malaria vector. The adult females can live up to a month (or more in captivity), but most probably do not live more than two weeks in nature. Based on the intrinsic incubation period of malaria parasites, ranging from the entry of the parasite into the human body up to the onset of clinical symptoms, *Plasmodium falciparum* takes 8-11 days, while *Plasmodium vivax* takes 12-17 days [3], [6], [11].

From those rational, it can be assumed that since the *Anopheles* laid their eggs, they still need approximately a month before the offspring are mature enough to fly and spread malaria. Therefore, the weather condition that provide the condition for *Anopheles* to breed may not affect directly to the malaria transmission and malaria incidence in the next month, but it may also affect to the next two or three month. Therefore, to cover all the correlation possibility, the system will designed to predict the malaria incidence for the 1st, 2nd, and 3rd month ahead.

The problem arise as the GMDH PNN network architecture only allows a single output, thus some modification in main process need to be done in order to implement this design process. The system then built to contain multiple learning processes that ran in parallel manner as many as the desired outputs. Each learning process would make a prediction model for each desired month (next 1st month, 2nd month, and 3rd month). The parallel prediction design is shown in Fig. 3
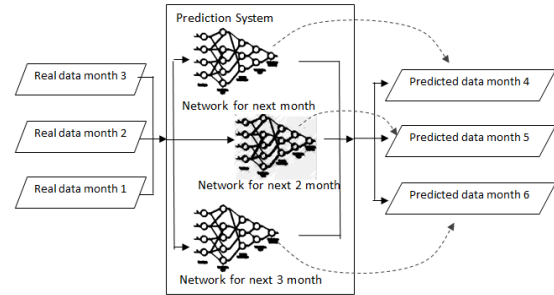

Fig. 3. Prediction design

V. MODIFIED GMDH PNN

In GMDH PNN, every time a new layer created the network will try every possible neuron with all possible input combination. Thus for a neuron that took 2 inputs from n available inputs, the combination that tried is C_n^2 combinations. To reduce the time and computation, the entropy and information gain for each input was calculated. Afterwards, rather than blindly try all possible combination, only a portion of best input combination based on the Information Gain that will be tried.

Entropy is a parameter to estimate heterogeneity from a set of data. The diversity in a set of data, determine higher the entropy value. Entropy is calculated by formula of Eq. (5) [12].

$$Gain(S, A) = Entropy(S) - \sum_{veValues(A)} \frac{|S_v|}{S} Entropy(S_v) \quad (5)$$

where:

S = attribute,

c = number of class,

p_i = number of sample in class i

Information Gain is a parameter to estimate the effectiveness of an attribute to classify the target. Information Gain of an attribute A is written is Eq. (6) [12].

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{S} Entropy(S_v) \quad (6)$$

where:

A = attribute
V = possible value in A
Values (A) = set of possible values in A
$|S_v|$ = number of sample for value v
$|S|$ = number of sample data,
Entropy (S_v) = entropy for samples that contain value v

By adding information gain, each input attribute will be graded by their effectiveness to target. Then, only combinations of the most effective attributes will be tried as input neuron in each layer. This will make the training process runs faster as not all input combination will be tried but still maintaining the accuracy as the selected attributes are those that have high effectiveness to target

VI. EXPERIMENT AND ANALYSIS

Based on [3], the growth of mosquitoes depends on some weather factors, which are temperature, rainfall, humidity, and length of daylight. However, based on the Health Department of Ministry of Indonesia, the weather factors that proposed only consist of temperature, rainfall excess, and precipitation day. Thus, in this thesis, scenario 1 and scenario 2 was proposed.

The first scenario is conducted to investigate the influence of three weather factors proposed by Health Department of Ministry of Indonesia to the performance of malaria prediction. The second scenario was aimed to investigate the influence of seven weather factors to the performance of malaria prediction compared to the first scenario

To improve the performance of the prediction system, the modified PNN was designed. To examine the effect of the modification, scenario 3 was proposed. . This scenario was aimed to investigate the influence of the modification to the performance of the system.

In previous work, malaria incidence typically conducted using Autoregressive and Neural Network. In order to compare the result of prediction model with scenario 2, the scenario 4 and scenario 5 was proposed. The fourth and fifth scenarios was aimed to investigate the performance of Feed-forward ANN and ARIMA in predicting malaria incidence using the same input and output properties with other scenario. The network architecture in scenario 4 was defined using 25 hidden layers and 15 neurons in each hidden layer. In scenario 5, the parameter used in ARIMA(d,p,q) is the combination of integration (d), autoregressive (p) and moving average (q) of ARIMA(1,1,0).

The input series combination examined in every scenario is from 1 to 5 previous months. The PNN algorithm is based on the GMDH method and utilizes a class of polynomials. By choosing the most significant input variables and polynomial order among these various types of forms available, we can obtain the best of the extracted partial descriptions according to both selecting nodes of each layer and generating additional layers until the best performance is reached [8], [13]. Polynomial function combination is examined using only bilinear (Eq. 7), biquadratic (Eq. 8), and triquadratic (Eq. 9).

$$Bilinear \quad p_0 + p_1 x_1 + p_2 x_2 \quad (7)$$

$$Biquadratic \quad p_0 + p_1 x_1 + p_2 x_2 + p_3 x_1 x_2 + p_4 x_1^2 + p_5 x_2^2 \quad (8)$$

$$Triquadratic \quad p_0 + p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_1 x_2 +$$
$$p_5 x_2 x_3 + p_6 x_1 x_3 + p_7 x_1^2 + p_8 x_2^2 + p_9 x_3^2 \quad (9)$$

Criterion function is the function that determines the quality of the network. While the criterion of the latest layer is still better than the previous layer, than the training process continued. The criterion function examined are criterion using RMSE validation data, using RMSE validation and training data, Corrected Akaike's Information Criterion (AICC), and using Minimum Description Length (MDL). From the combination of 3 polynomial functions and 4 criterion function there are 12 experiments in total. The experiments are listed in Table I

TABLE I. EXPERIMENT COMBINATION

No Exp	Neuron Function	Criterion Function	No Exp	Neuron Function	Criterion Function
1	biquadratic	RMSE Training	7	triquadratic	AICC
2	biquadratic	RMSE Training + Validation	8	triquadratic	MDL
3	biquadratic	AICC	9	bilinear	RMSE Training
4	biquadratic	MDL	10	bilinear	RMSE Training + Validation
5	triquadratic	RMSE Training	11	bilinear	AICC
6	triquadratic	RMSE Training + Validation	12	bilinear	MDL

Every experiment will be examined using 1 to 5 input series to predict 1st, 2nd, and 3rd month prediction. Therefore, there are 60 examinations that will be conducted for every month model prediction. Each examination will produce one prediction model.

Each combination in every scenario was examined and all the result calculated in Mean Absolute Percentage Error (MAPE). The MAPE calculation was used to determine which combination that performed a better result.

Because of some limitation in computing process, some prediction model trained using GMDH PNN resulting error at was not able to predict the output from some input series. This particular error existed because the polynomial functions reached the maximum value that was calculated by the computer environment. This error did not result the false prediction but due to limitation. Thus, to determine the best model, the quality of each model is calculated to minimize MAPE and failed prediction rate using optimization function. This function is based on Fitness function from Genetic Algorithm with minimizing [13].

$$Quality_i = \frac{1}{(MAPE_i + \%FAIL_i + k)} \qquad (10)$$

where

$Quality_i$: quality of i-th data
$MAPE_i$: quality of i-th data
$\%Fail_i$: quality of i-th data

k : constant small value (k = 0.1)

The function will maximize the minimizing problem as solved by Eq. (10). The higher the value of quality indicates the better performance of the system. The highest value of quality will be achieve when the MAPE is resulting 0 (no error prediction) and %Fail is also 0 (no failed prediction) which is Quality = 10.

TABLE II: BEST MODEL COMPARISON FROM ALL SCENARIOS

Month Prediction	Scenario	Function	Criterion	Num Series	Accuracy	Quality	Time	
							Training	Testing
1	1	Triquadratic	RMSE Train	1	73.68	0,038	0,334	0,072
	2	Biquadratic	AICC	5	80.97	0,052	1,778	0,083
	3	Biquadratic	RMSE Train+Val	5	88.02	0,083	16,848	0,278
	4			4	69.63	0,033	168,727	0,020
	5			5	65.93	0,029	6,671	0,333
2	1	Triquadratic	RMSE Train	2	82.86	0,058	3,687	0,324
	2	Bilinear	RMSE Train+Val	5	80.75	0,052	70,418	2,753
	3	Biquadratic	RMSE Train+Val	5	86.12	0,072	10,323	0,761
	4			4	69.63	0,033	168,727	0,020
	5			5	66.12	0,029	6,160	0,334
3	1	Triquadratic	AICC	5	72.37	0,036	2,372	0,022
	2	Bilinear	RMSE Train+Val	5	78.93	0,047	37,304	3,170
	3	Biquadratic	RMSE Train+Val	4	83.90	0,062	2,000	0,256
	4			4	76.13	0,042	151,176	0,020
	5			5	71.13	0,035	6,101	0,317

The malaria incidence forecasting system correlating the incidence to weather pattern performed a good performance with accuracy of 88.02%, 86.12%, and 83.90% for 1st month, 2nd month, and 3rd month prediction as shown in Table II. The average accuracy of prediction model trained using only three weather factors as input was increased by 9.87%, 10.54%, and 11.55% when the prediction model trained using seven weather factor as input. This indicates that with more weather factors the prediction model was able to perform better generalization.

The performance of GMDH PNN in predicting malaria overpowers the performance of prediction using ANN BP with the difference of average accuracy by 23.18%, 21.71%, and 24.68%. Although the learning time to build prediction model using ARIMA was the fastest, the performance of ARIMA still overpowered by GMDH PNN with the difference of average accuracy by 19.81%, 19.45%, and 14.93%. The result showed that prediction trained using 5 input series perform a better prediction than using less input series.

The modification made to GMDH Polynomial Neural Network by reducing the calculating complexity and optimizing the neuron selecting process was able to improve the prediction model by reducing the learning time by 66%, 78%, and 73% for 1st, 2nd, and 3rd month while maintaining its high accuracy. The average accuracy of prediction using the original GMDH PNN was increased by 1.3%, 0.66%, and 1.42% when the prediction model trained using the modified GMDH PNN. The prediction result using modified GMDH PNN is shown in Fig. 4.

VII. CONCLUSION

Modified GMDH Polynomial Neural Network, with addition of Entropy and Information Gain to choose the input combination for each neuron in every layer, results a better prediction model than the original GMDH PNN. The modified method was also able to speed the training process by reducing the computational process.

By some observation, the optimal parameter input series were selected for the optimal training process of GMDH PNN. They were 4 to 5 input series using all weather factors data, while trained using polynomial function of Biquadratic and Triquadratic. The optimal criterion function observed was using RMSE Train+Val and using Corrected Akaike's Information Criterion.

The prediction model trained with more weather factors as input performs better than prediction model trained using only three weather factors with accuracy improvement of 9.87% for 1st month prediction, 10.54% for 2nd month prediction and 11.55% for 3rd month prediction. The performance of GMDH PNN overpowered the performance of prediction model trained using ANN BP and ARIMA in term of accuracy.

The implementation of Modified GMDH PNN was able to improve the system performance based on the accuracy rates and quality. The modified GMDH PNN were able to reach prediction accuracy of 88.02% with quality of 0.083 for 1st month prediction, 86.12% with quality of 0.072 for 2nd

month prediction, and 83.90% with quality of 0.062 for 3rd month prediction and also reduce the average training time by 72% compared to the original GMDH PNN.

Most of failed predictions were the result of the limited calculation ability of the environment. The system was able to be improved by optimizing the data selection and normalization process.
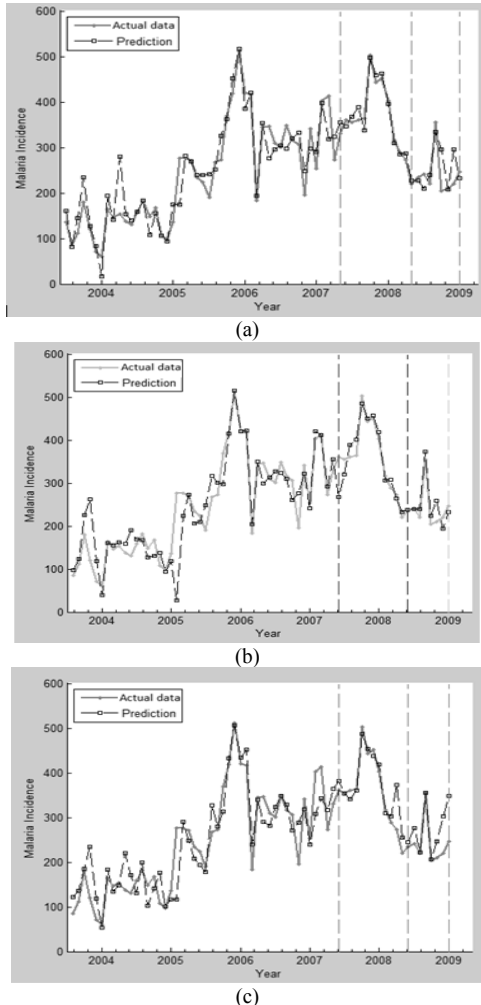

(a)

(b)

(c)

Fig. 4. Prediction result using modified GMDH PNN
(a) 1st month prediciton, (b) 2nd month prediction, (c) 3rd month prediction

VIII. FUTURE WORK

The treatment method in data preparation will affect the performance of the learning process. Any other normalization and data complexity reduction such as PCA can be applied to improve the prediction performance.

Criterion function is the function that determines the quality of the network and the one that decides whether the learning process needs to be continued or not. Any other quality measurement function and technique may help the learning process to make a better prediction model. Study to minimalize the failed prediction is needed so the performance of GMDH PNN will fully visible.

More data analysis about the connection from weather data and other attributes with malaria incidence was needed to select the data attribute that has the most effect and correlation to the malaria incidence. More additional attributes that may have some contribute to malaria incidence such as NDVI data might improve the prediction result.

REFERENCES

[1] A. G. Elipe, A. Otero, M. V. Herp, and A. A. Jaime, "Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997–2003," *Malaria Journal*, vol. 6, no. 1, pp. 129, September 2007.

[2] WHO. (2012). Indonesia-Malaria Situation in SEAR Countries. [Online]. Available: www.searo.who.int/en/Section10/Section21/Section340_4022.htm

[3] S. U. K. Hadi, S. H. Sigit, and S. Sukowatu, "Hubungan Iklim, Kepadatan Nyamuk Anopheles dan Kejadian Penyakit Malaria," *J. Entomol. Indon*, vol. 7, no. 1, pp. 42-53, April 2010.

[4] P. Liatsis, A. Foka, J. Y. Goulermas, and L. Mandic, "Adaptive polynomial neural networks for times series forecasting," in *Proc.49th International Symposium ELMAR*, Zadar, Croatia, 2007, pp. 35 - 39.

[5] A. G. Ivakhnenko, "The Review of Problems Solved by Algorithms of the Group Method Data Handling," *Pattern Recognition and Image Analysis*, vol. 5, no. 4, pp. 527-535, 1995.

[6] E. Loha and B. Lindtjørn, "Model variations in predicting incidence of Plasmodium falciparum malaria using 1998-2007 morbidity and meteorological data from south Ethiopia," *Malaria Journal*, vol. 9, pp. 166, 2010.

[7] R. Ghazali, A. J. Hussain, P. Liatsis, and H. Tawfik, "The application of ridge polynomial neural network to multi-step ahead financial time series prediction," *Neural Comput & Applic*, vol. 17, pp. 311-33, July 2007.

[8] S. K. Oh and W. Pedrycz, "The design of self-organizing Polynomial Neural Networks," *Information Sciences*, vol. 141, pp. 237–258, 2002.

[9] F. Lemke and J. A. Müller, *Self-Organizing Data Mining Based on GMDH Principle*.

[10] J. Peralta, G. Gutierrez, and A. Sanchis, "Design of Artificial Neural Networks Based on Genetic Algorithms to Forecast Time Series," *Innovations in Hybrid Intelligent Systems*, Springer, vol. 44, pp. 231-238, 2007.

[11] B. Marston, *The Natural History of Mosquitoes*, New York: The Mac Mollon Co, 1949.

[12] Suyanto, *Artificial Intelligence: Searching, Reasoning, Planning, and Learning, Bandung*, Penerbit Informatics, 2007.

[13] S. Farzi, "The Design of Self-Organizing Evolved Polynomial Neural Network Based on Learnable Evolution Model 3," *The International Arab Journal of Information Technology*, vol. 9, no. 2, pp. 124-132, 2012.

WHO. (2011, September) WHO Regional Office for South-East Asia. [Online]. Available: www.searo.who.int/en/Section10/Section21/Section340_4018.htm

Anditya Arifianto was 1989 in Jember, East Java. He graduated from Institut Teknologi Telkom with bachelor's degree in Computer Science in 2011 and master's degree in Informatics Engineering in 2013.

He is currently teaching Computer Science at the Faculty of Informatics IT Telkom. Research interest on the development of Artificial Intelligence elaborated case study Image Processing, Time Series Prediction, and Text and Data Mining

Ari Moesriami Barmawi received her bachelor degree from Bandung Institute of Technology in 1985. She received both MSc and PhD degree in computer science from Keio University, Japan in 1997 and 2001 respectively. She is currently a lecturer and the Dean of Telkom Institute of Technology Graduate School, Indonesia. Her main research activities concern the intelligent systems and data security.

Agung Toto Wibowo was born in Kudus city in 1981, getting bachelor's degree in Computer Science at STT Telkom in 2005, also master's degree in Electrical Engineering at Institut Teknologi Bandung in 2009.

He is currently teaching Computer Science at the Faculty of Informatics IT Telkom since 2006. Research interest on the development of Artificial Intelligence elaborated case study Plagiarism detection, Time Series Prediction and Cutting stock problem