Ramadhan Short-Term Electric Load: A Hybrid Model of Cycle Spinning Wavelet and Group Method Data Handling (CSW-GMDH)

Rezzy Eko Caraka, Rung Ching Chen*, Toni Toharudin, Bens Pardamean, Sakhinah Abu Bakar, Hasbi Yasin

Abstract—**In general, performing a nonlinearity time series analysis in the modeling of data can reach a robust and increase the quality of the results. Wavelet methods have successfully been applied in a great variety of applications for modeling also forecasting. Wavelet Transform divided into two categories. There is continuous wavelet (CWT) and a discrete wavelet transform (DWT). Cycle spinning unlike the discrete wavelet transform (DWT), is highly redundant, non-orthogonal, also defined naturally for all sample sizes. There is a Group Method of Data Handling (GMDH) algorithm, which is a multivariate analysis method can be used in modeling and identifying uncertainty on linear also nonlinearity systems. In this paper, we aim to explain the combination of À-Trous wavelet transforms applied on cycle spinning and group method of data handling (GMDH) in data of short-term electric load holy month of Ramadhan from 2014 to 2015**

Index Terms—**Wavelet, GMDH, Electrical Load, Time series**

I. Introduction

Cycle Spinning Wavelet and Group Method Data Handling (CSW-GMDH) was obtained by combining two are cycle spinning wavelet DWT and GMDH. The DWT was applied to the input variable where the DWT decomposed each of the variables into some components. The DWT decomposed the variable using Mallat DWT algorithm [1]. The Mallat algorithm translates each data without losing the information about the element in the original data [2].

Wavelet used in wide area, Dai (2013) using wavelets and grey correlation analysis combined to assist the improved GM(1,1) model to predict HRG's lifetime [3]. Chew (2009) presents an image compression technique called the Tuned Degree-K Zerotree Wavelet (TDKZW) coding which is targeted at single-encoding, multiple-decoding image processing applications. In the proposed work, the degree of zerotree tested is tuned in each encoding pass to achieve an optimal compression performance [4].

Wavelet based model makes time series predictions based on dependencies of series to the historical information. The data pattern will be captured well when the series have high dependencies to the historical information [5].

The latest development related with wavelet analysis and Neural Network (NN) for time series forecasting shows that some researches associated with combination of wavelet and neural networks resulting a hybrid model which is known as Wavelet Neural Network (WNN) are begun by some wavelet and NN researchers [6].

The disadvantage of WNN is the lack of structured method to determine the optimum level of WNN factors, which are mostly set by trial and error. The factors affecting the performance of WNN are the level of MODWT decomposition, the wavelet family, the lag inputs, and the number of neurons in the hidden layer [7]. However, MODWT decomposition can be combine with mRMR for constructing forecasting model for time series and model resulted just contained coefficients that were considered important enough to give influence to the present value [8].

The decomposition process is a continuous filtering process where the input variable decomposed into an approximation and detail components [9]. It allows the GMDH to learn about the characteristics of the data and produce good estimation. The CSW-MDH model is the improvement of GMDH model by combining the two methods which are the cycle spinning (CS) model and the Group Method of Data Handling (GMDH). In this study, multivariate input variables were used. DWT decomposed the input variables into several components that are approximate and details. The numbers of detail components are dependent on the number of resolution level implemented for that DWT. The number of resolution levels achieved in this study is 2, 3, and 4. After the original input variables are decomposed using DWT, then the useful components of approximate and details are chosen using correlation. Only

Manuscript received April 04, 2019; revised July 17, 2019. This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST-107-2221-E-324-018-MY2, partial supported by BDSRC Bina Nusantara University, and Department of Statistics Padjadjaran University.

R.E.Caraka. College of Informatics, Chaoyang University of Technology, Taichung City, 41349, Taiwan (R.O.C.), and Department of Statistics, Padjadjaran University, Bandung, Indonesia

*R.C. Chen. College of Informatics, Chaoyang University of Technology, Taichung City, 41349, Taiwan (R.O.C.); (e-mail: crching@cyut.edu.tw)

T.Toharudin, Department of Statistics, Padjadjaran University, Bandung, Indonesia

B.Pardamean, Bioinformatics Data Science Research Center, Bina Nusantara University, and BINUS Graduate Program – Master of Computer Science, Bina Nusantara University.

S.A.Bakar, School of Mathematical Sciences, FST, the National University of Malaysia.

H,Yasin, Department of Statistics, Diponegoro University, Semarang, Indonesia.

the active components after decomposition. The two primary goals of electric load time series analysis are to (i) determine the nature of the phenomenon represented by the sequence of observations electric load and possibly provide a model for those observations[10]. (ii) Forecast future possible electric load values[11].

Methods of time series analysis can be performed in either the frequency or time domain, and there is no single methodology to analyses a time series data that is suitable for all fields. Usually, the field from which the data was obtained and aim of the study will dictate the appropriate methodology [12]. In this paper, we employ the wavelet transform as a tool to smooth and minimize the noise present in the Electric load; since the noise contained in the series distorts the estimated parameters. Besides, the smoothing effect on the forecasting of the electric load while using GMDH.

II. RESEARCH METHOD

A. Wavelet

Wavelet is a name for the small wave up and down in a specified period[13]. Meanwhile, as in comparison to the big waves, an, e.g. sine wave function that down also moves up on a plot of *sin* (u) $\in(-\infty,\infty)$. Wavelet functions divided into two functions, namely wavelet father *(ϕ)* and wavelet mother *(ψ)*.Wavelet function has the properties:

$$\int_{-\infty}^{\infty}\emptyset(x)dx = 1 \text{ and } \int_{-\infty}^{\infty}\psi(x)dx = 0 \quad (1)$$

With dyadic dilation and translation integer, wavelet father and wavelet mother shows scale function can be written:

$$\emptyset_{j,k}(x) = 2^{\frac{j}{2}}\emptyset(2^j x - k) \text{ And } \psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k) \quad (2)$$

Function $\emptyset_{j,k}(x)$ and $\psi_{j,k}(x)$Which is orthogonal to have properties

$$\int_{-\infty}^{\infty}\emptyset_{j,k}(x)\,\emptyset_{j,k'}(x)dx = \delta_{k,k'}, \quad (3)$$

$$\int_{-\infty}^{\infty}\psi_{j,k}(x)\,\emptyset_{j,k'}(x)dx = 0, \quad (4)$$

$$\int_{-\infty}^{\infty}\psi_{j,k}(x)\psi_{j',k'}(x)dx = \delta_{j,j'},\delta_{k,k'}, \quad (5)$$

With $\delta_{i,j} = \begin{cases}1 \; if \; i = j \\ 0 \; if \; 1 \neq j\end{cases}$

Haar wavelets are based on the functions which were introduced by Hungarian mathematician Alfred Haar in 1910 [14]. The Haar wavelets are made up of piecewise constant functions and are mathematically the simplest among all the wavelet families. A useful feature of these wavelets is the possibility to integrate them analytically Arbitrary times [14]. They can be interpreted as a first order Daubechies wavelet[15]. The Haar wavelets have been applied for solving several problems of mathematical calculus.

$$\psi(x) = \begin{cases}1, & 0 \le x < \frac{1}{2} \\ -1, & \frac{1}{2} \le x < 1 \\ 0, & x \text{ otherwise}\end{cases}$$

And

$$\varphi(x) = \begin{cases}1, & 0 \le x < 1 \\ 0, & x \text{ otherwise}\end{cases}$$

Table I represents the wavelet and indicates the first letter of the name, *d* for the doublet, *s* for symmlet and *c* for coiflet. Numbers indicate the length of wavelet support (width) and smoothness [16]. Wavelet with large numbers such as *D20* or *C30* is relatively wide and smooth [9]. The following Table I of four types of orthogonal wavelet [17].

TABLE I
WAVELET ORTHOGONAL

Type	Wavelet
Haar	"haar" or "d2"
Daublets	"d4" "d6" "d8" "d12" "d14" "d16" "d18" "d20"
Symmlet	"s4" "s6" "s8" "s12" "s14" "s16" "s18" "s20"
Coiflet	"c6" "c12" "c24" "c30"

B. Wavelet Transform

Wavelet function can form a base in the $L^2(R)$ with $L^2(R) = \{f \mid \int_{-\infty}^{\infty} f^2(x)dx < \infty\}$. As a result every $f \in L^2(R)$ can be declared as a combination in linear a base that was built by wavelet

$$f(x) = \sum_{k\in Z} c_{J,k}\,\phi_{J,k}(x) + \sum_{j<J}\sum_{k\in Z} d_{j,k}\psi_{j,k}(x) \quad (6)$$

with :

$$c_{J,k} = \langle f, \phi_{J,k}\rangle = \int_{-\infty}^{\infty} f(x)\phi_{J,k}(x)dx \quad (7)$$

$$d_{j,k} = \langle f, \psi_{j,k}\rangle = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx \quad (8)$$

Function *f* produce series is not up to, but function f can be accessed with the final results were entered into a limited to good use with the index *J*, with *A* so it can be declared as final results were entered into a becomes component's and components detail D. Using wavelet shrinkage [18] to smooth data can sometimes lead to visual artifacts which, are caused by features either in the signal or basis or both. The resulting smoothed data can exhibit Gibbs phenomena, and alternate under an overestimation of a specific target. The size of the effect is directly related to the location, for example, due to the nature of the Haar wavelet, an effect in or very close to a dyadic location will not cause any problems. However, at other locations, this can cause quite a significant innocence of the result. One method to mitigate these issues is to use cycle spinning [19]. It works by:

- Shifting the data, a random number of positions.
- Curve estimation of the shifted data.
- Un-shifting the estimate to their original positions.

Cycle Spinning, also related to notions of maximal overlap, wavelet frames, shift invariant, DWT wavelet frames, translation DWT is with the basic idea of Downsampling values removed from discrete wavelet transform. Cycle spinning unlike the conventional discrete wavelet transform (DWT), which is highly redundant also non-orthogonal. In other words, is defined naturally for all sample sizes, *N*. To give an example there is a data of time series *X*, then wavelet transform will produce vector column $W_1, W_2, \dots, W_{J_0}$ and $V_{J_0}$with each measuring$N$. Smoothing Coefficient that comes from the data *X* came from multiple repeatedly from *X* with filter scale ($\tilde{g}$) and ($\tilde{g}$) wavelet filter($\tilde{h}$). This sensitivity is down sampling of wavelet filter output and filter scale at each stage of the algorithm pyramid.

$$\widetilde{\mathcal{W}}_{1,t} \equiv \sum_{l=0}^{L-1} \tilde{h}_l X_{t-l \bmod N} \qquad (9)$$

t = 0, gained: $\widetilde{\mathcal{W}}_{1,0} \equiv \widetilde{\mathcal{W}}_0^T X = \sum_{l=0}^{N-1} \tilde{h}^{\circ}_{-l \bmod N} X_l$,
so, $\widetilde{\mathcal{W}}_0^T = [\tilde{h}^{\circ}_0, \tilde{h}^{\circ}_{N-1}, \tilde{h}^{\circ}_{N-2}, \dots, \tilde{h}^{\circ}_1]$
$L \leq N$, Filter periodic matter takes the form simple

$$\tilde{h}^{\circ}_l = \begin{cases} \tilde{h}_l, & 0 \leq l \leq L-1; \\ 0, & L \leq l \leq N-1 \end{cases}$$

So the line of the first $\widetilde{\mathcal{W}}_1$ is

$$\widetilde{\mathcal{W}}_0^T = [\tilde{h}_0, 0, \dots, 0, \tilde{h}_{L-1}, \dots, \tilde{h}_1]$$

Because $L = 4$, thus $\tilde{h}_{L-1} = \tilde{h}_3$ and $\tilde{h}_L$ high-zero until $\tilde{h}_{N-1}$, with the number of elements that high-zero is $N - L$. So the first line matrix $\widetilde{\mathcal{W}}_1$

$$[\tilde{h}_0 \quad 0_3 \quad 0_3 \quad \dots_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad \tilde{h}_3 \quad \tilde{h}_2 \quad \tilde{h}_1]$$

Second Period when line $\widetilde{\mathcal{W}}_1$ when $t = 1$, gained

$$\widetilde{\mathcal{W}}_{1,0} \equiv \widetilde{\mathcal{W}}_1^T X = \sum_{l=0}^{N-1} \tilde{h}^{\circ}_{1-l \bmod N} X_l,$$
$$\widetilde{\mathcal{W}}_1^T = [\tilde{h}^{\circ}_1, \tilde{h}^{\circ}_0, \tilde{h}^{\circ}_{N-1}, \dots, \tilde{h}^{\circ}_2]$$

The second line from $\widetilde{\mathcal{W}}_1$ is

$$\widetilde{\mathcal{W}}_1^T = [\tilde{h}_1, \tilde{h}_0, 0, \dots, 0, \tilde{h}_{L-1}, \dots, \tilde{h}_2]$$
$$[\tilde{h}_1 \quad \tilde{h}_0 \quad 0_3 \quad \dots_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad 0_3 \quad \tilde{h}_3 \quad \tilde{h}_2]$$

Applied goes on to $t = N - 1$, and matrix filter wavelet with the structure As follows:

$$\widetilde{\mathcal{W}}_1 = \begin{bmatrix} \tilde{h}_0 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 \\ \tilde{h}_1 & \tilde{h}_0 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & \tilde{h}_2 \\ \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 \\ \vdots_3 & \vdots_3 & \vdots_3 & \dots_3 & \vdots & \vdots & \vdots & \vdots_3 & \vdots_{33} & \vdots_{33} & \vdots_3 & \vdots_3 \\ 0_3 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & 0_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & 0_3 \\ 0_3 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 \end{bmatrix}$$

$\widetilde{\mathcal{V}}_1$ drawn up $\widetilde{\mathcal{V}}_{1,t} \equiv \sum_{l=0}^{L-1} \tilde{g}_l X_{t-l \bmod N}$, $\tilde{h}_l$ will be changed to $\tilde{g}_l$.

$$\widetilde{\mathcal{V}}_1 = \begin{bmatrix} \tilde{g}_0 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{g}_3 & \tilde{g}_2 & \tilde{g}_1 \\ \tilde{g}_1 & \tilde{g}_0 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{g}_3 & \tilde{g}_2 \\ \tilde{g}_2 & \tilde{g}_1 & \tilde{g}_0 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{g}_3 \\ \vdots_3 & \vdots_3 & \vdots_3 & \dots_3 & \vdots & \vdots & \vdots & \vdots_3 & \vdots_{33} & \vdots_{33} & \vdots_3 & \vdots_3 \\ 0_3 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & \tilde{g}_3 & \tilde{g}_2 & \tilde{g}_1 & \tilde{g}_0 & 0_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & \tilde{g}_3 & \tilde{g}_2 & \tilde{g}_1 & \tilde{g}_0 & 0_3 \\ 0_3 & 0_3 & 0_3 & \dots_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{g}_3 & \tilde{g}_2 & \tilde{g}_1 & \tilde{g}_0 \end{bmatrix}$$

It can be written into the equation:

$$\begin{bmatrix} \widetilde{\boldsymbol{W}}_1 \\ \widetilde{\boldsymbol{V}}_1 \end{bmatrix} = \begin{bmatrix} \widetilde{\mathcal{W}}_1 \\ \widetilde{\mathcal{V}}_1 \end{bmatrix} x = \widetilde{\mathcal{P}}_1 \boldsymbol{X}, \text{ With } \widetilde{\mathcal{P}}_1 = \begin{bmatrix} \widetilde{\mathcal{W}}_1 \\ \widetilde{\mathcal{V}}_1 \end{bmatrix}$$

and P_1^T is orthonomal matrix. Thus, for data reconstruction X, form coefficient will be done in the first level

$$\boldsymbol{X} = \widetilde{\mathcal{P}}_1^{-1} \begin{bmatrix} \widetilde{\boldsymbol{W}}_1 \\ \widetilde{\boldsymbol{V}}_1 \end{bmatrix} = \widetilde{\mathcal{P}}_1^T \begin{bmatrix} \widetilde{\boldsymbol{W}}_1 \\ \widetilde{\boldsymbol{V}}_1 \end{bmatrix} = \begin{bmatrix} \widetilde{\mathcal{W}}_1 \\ \widetilde{\mathcal{V}}_1 \end{bmatrix}^T \begin{bmatrix} \widetilde{\boldsymbol{W}}_1 \\ \widetilde{\boldsymbol{V}}_1 \end{bmatrix}$$

P orthogonal matrix, then $\widetilde{\mathcal{P}}_1^{-1} = \widetilde{\mathcal{P}}_1^T$. Second level of matrix $\widetilde{\mathcal{W}}_2$ With size NxN (applied also to matrix $\widetilde{\mathcal{V}}_2$ with replacing $\tilde{h}$ with $\tilde{g}$)

$$\widetilde{\mathcal{W}}_2 = \begin{bmatrix} \tilde{h}_0 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & 0_3 & \tilde{h}_2 & 0_3 & \tilde{h}_1 & 0_3 \\ 0_3 & \tilde{h}_0 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & 0_3 & \tilde{h}_2 & 0_3 & \tilde{h}_1 \\ \tilde{h}_1 & 0_1 & \tilde{h}_0 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_2 & 0_3 \\ \vdots_3 & \vdots_3 & \vdots_3 & \vdots_3 & \vdots & \vdots & \vdots & \vdots_3 & \vdots_{33} & \vdots_{33} & \vdots_3 & \vdots_3 \\ 0_3 & 0_3 & 0_3 & \tilde{h}_3 & 0_3 & \tilde{h}_2 & 0_3 & \tilde{h}_1 & 0_3 & \tilde{h}_0 & 0_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & 0_3 & \tilde{h}_2 & 0_3 & \tilde{h}_1 & 0_3 & \tilde{h}_0 & 0_3 \\ 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \tilde{h}_3 & 0_3 & \tilde{h}_2 & 0_3 & \tilde{h}_1 & 0_3 & \tilde{h}_0 \end{bmatrix}$$

The DWT filter must meet the following equation:

$$\sum_{l=0}^{L-1} h_l = 0, \sum_{l=0}^{L-1} h_l^2 = 1 \text{ And } \sum_{l=0}^{L-1} h_l h_{l+2n} = 0$$

Moreover, the scale filter must meet the following formula:

$$\sum_{l=0}^{L-1} g_l = \sqrt{2} \text{ or } \sum_{l=0}^{L-1} g_l = -\sqrt{2}, \ \sum_{l=0}^{L-1} g_l^2 = 1 \text{ And } \sum_{l=0}^{L-1} g_l g_{l+2n} = 0$$

For all integers, n is not zero. Table II explain the relationship wavelet filter and filter scale can be defined wavelet filter $\{\tilde{h}_l\}$ formed from $\tilde{h}_l \equiv h_l/\sqrt{2}$ and filter scale $\{\tilde{g}_l\}$ formed from $\tilde{g}_l \equiv g_l/\sqrt{2}$. So that the terms of a wavelet filter must meet the following equation:

$$\sum_{l=0}^{L-1} \tilde{h}_l = 0, \sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2} \text{ and } \sum_{l=0}^{L-1} \tilde{h}_l \tilde{h}_{l+2m} = 0$$

Filter scale must meet:

$$\sum_{l=0}^{L-1} \tilde{g}_l = 1, \sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2} \text{ And } \sum_{l=0}^{L-1} \tilde{g}_l \tilde{g}_{l+2m} = 0 \text{ where } m = 1, 2, \dots, (L/2) - 1.$$

Each length filter $L_j \equiv (2^j - 1)(L - 1) + 1$.

TABLE II
LENGTH FILTER

Level	filter ($L_j \equiv (2^j - 1)(L - 1) + 1$)	
	Haar	Daubechies
1	2	4
2	4	10
. . .	. . .	. . .
J_0	$(2^{j_0} - 1) + 1$	$3(2^{j_0} - 1) + 1$

Haar wavelet level $j = 1$ has $L_j = L = 2$. Can be written:
$\sum_{l=0}^{L-1} \tilde{h}_l = 0 \rightarrow \tilde{h}_0 + \tilde{h}_1 = 0$, $\sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2} \rightarrow \tilde{h}_0^2 + \tilde{h}_1^2 = \frac{1}{2}$
And $\sum_{l=0}^{L-1} \tilde{h}_l \tilde{h}_{l+2m} = 0 \rightarrow \tilde{h}_0 \tilde{h}_2 + \tilde{h}_1 \tilde{h}_3 = 0$
$L_j = L = 2$ and $\{\tilde{h}_l : l = 0, \dots, L-1\}$, then $\tilde{h}_l : \{\tilde{h}_0, \tilde{h}_1\}$ $\tilde{h}_2$ and $\tilde{h}_3$ Is considered zero

$$\sum_{l=0}^{L-1} \tilde{h}_l = 0 \rightarrow \tilde{h}_0 + \tilde{h}_1 = 0$$
$$\tilde{h}_1 = -\tilde{h}_0$$

Then the equation is substituted, it is obtained:

$$\sum_{l=0}^{L-1} \tilde{h}_l^2 = \frac{1}{2} \rightarrow \tilde{h}_0^2 + \tilde{h}_1^2 = \frac{1}{2}$$
$$\tilde{h}_0^2 + (-\tilde{h}_0)^2 = \frac{1}{2}$$
$$2\tilde{h}_0^2 = \frac{1}{2}$$
$$\tilde{h}_0 = \sqrt{\frac{1}{4}} = \frac{1}{2}$$

and

$\tilde{h}_1 = -\tilde{h}_0$; $\tilde{h}_1 = -\frac{1}{2}$

So the level wavelet filters $j = 1$ of Haar wavelet. $\{\tilde{h}_0 = \frac{1}{2}, \tilde{h}_1 = -\frac{1}{2}\}$. Haar wavelet filter scale for level $j = 1$ has $L_j = L = 2$ can be calculated as follows:

$$\sum_{l=0}^{L-1} \tilde{g}_l = 1, \sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2} \text{ And } \sum_{l=0}^{L-1} \tilde{g}_l \tilde{g}_{l+2m} = 0 \qquad (10)$$

$$\sum_{l=0}^{L-1} \tilde{g}_l = 1 \rightarrow \tilde{g}_0 + \tilde{g}_1 = 1, \sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2} \rightarrow \tilde{g}_0^2 + \tilde{g}_1^2 = \frac{1}{2}$$

And $\sum_{l=0}^{L-1} \tilde{g}_l \tilde{g}_{l+2m} = 0 \rightarrow \tilde{g}_0 \tilde{g}_2 + \tilde{g}_1 \tilde{g}_3 = 0$

$L_j = L = 2$ and $\{\tilde{g}_l : l = 0, \ldots, L-1\}$, Then $\tilde{g}_l : \{\tilde{g}_0, \tilde{g}_1\}$ For $\tilde{g}_2$ and $\tilde{g}_3$ considered zero
$\sum_{l=0}^{L-1} \tilde{g}_l = 0 \rightarrow \tilde{g}_0 + \tilde{g}_1 = 1$
$\tilde{g}_1 = 1 - \tilde{g}_0$

Then the equation is substituted, it is obtained:

$$\sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2} \rightarrow \tilde{g}_0^2 + \tilde{g}_1^2 = \frac{1}{2}$$
$$\tilde{g}_0^2 + 1 - 2\tilde{g}_0 + \tilde{g}_0^2 = \frac{1}{2}$$
$$2\tilde{g}_0^2 - 2\tilde{g}_0 + \frac{1}{2} = 0$$
$$\tilde{g}_0 = \frac{1}{2}; \tilde{g}_1 = 1 - \tilde{g}_0 = \frac{1}{2}$$

So the filter scale level j = 1 of the Haar wavelet can be describe$\{\tilde{g}_0 = \frac{1}{2}, \tilde{g}_1 = \frac{1}{2}\}$.

C. Pyramid Algorithm

In the orthonormal Discrete Wavelet Transform (DWT) the wavelet coefficients are related to non-overlapping differences of weighted averages from the original observations that are concentrated in space. The most common way to implement DWT is the pyramid algorithm [20]. More information on the variability of the signal could be obtained considering all possible differences at each scale [21]. Which is considering overlapping differences, and this is precisely what the maximal overlap algorithm does [22]. Indeed, the term maximal overlap refers to the fact that all reasonable shifted time intervals are computed. As a consequence, the orthogonality of the transform is lost, but the number of wavelets and scaling coefficients at every scale is the same as the number of observations. Thus, the maximal overlap DWT coefficients may be considered the result of a simple modification in the pyramid algorithm used in computing DWT coefficients through not downsampling the output at each scale and inserting zeros between coefficients in the wavelet and scaling filters. A circular filter$\{\tilde{h}_l =$ *0, ...,L−1,* with length*2j−1L−1+1* has a series

$$\tilde{h}_0, \underbrace{0, \ldots, 0}_{2^{j-1}\text{-1 zeroes}}, \tilde{h}_1, \underbrace{0, \ldots, 0}_{2^{j-1}\text{-1 zeroes}}, \ldots, \tilde{h}_{L-2}, \underbrace{0, \ldots, 0}_{2^{j-1}\text{-1 zeroes}}, \tilde{h}_{L-1}$$

Moreover, has a transfer function$\widetilde{H}(2^{j-1}f)$. The elements of $\{\widetilde{W}_{j,t}\}$ is obtained from $\{\tilde{V}_{j-1,t}\}$ by the formula,
$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L-1} \tilde{h}_l \tilde{V}_{j-1,t-2^{j-1}l \bmod N}, \; t = 0,1,\ldots,N-1$
With the same description, then:
$\tilde{V}_{j,t} \equiv \sum_{l=0}^{L-1} \tilde{g}_l \tilde{V}_{j-1,t-2^{j-1}l \bmod N}, \; t = 0,1,\ldots,N-1$

Both of these equations are the basis for cycle spinning pyramid algorithm. If defined $\tilde{V}_{0,t} = X_t$,then the above equation produces coefficients of scale wavelet and coefficients of the first level.

$$\tilde{V}_{0,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{W}_{1,t+2^{1-1}l \bmod N} + \sum_{l=0}^{L-1} \tilde{g}_l \tilde{V}_{1,t+2^{1-1}l \bmod N} \quad (11)$$
$$= \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{W}_{1,t+1l \bmod N} + \sum_{l=0}^{L-1} \tilde{g}_l \tilde{V}_{1,t+1l \bmod N}$$
$$= \sum_{l=0}^{N-1} \tilde{h}_l \widetilde{W}_{1,t+1l \bmod N} + \sum_{l=0}^{N-1} \tilde{g}_l \tilde{V}_{1,t+1l \bmod N}$$
$$= \widetilde{\mathcal{W}}_1^T \widetilde{\boldsymbol{W}}_j + \tilde{\mathcal{V}}_1^T \widetilde{\boldsymbol{V}}_j$$
$$= \widetilde{\mathcal{D}}_1 + \tilde{\mathcal{S}}_1$$

If $\widetilde{\boldsymbol{V}}_{\boldsymbol{0}} = \boldsymbol{X}$, then the application repeatedly subjected to the above equation to J_0,level l, then its matrix notation can be written

$$\boldsymbol{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{W}}_j^T \widetilde{\boldsymbol{W}}_j + \tilde{\mathcal{V}}_{J_0}^T \widetilde{\boldsymbol{V}}_{J_0}$$
$$\boldsymbol{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \tilde{\mathcal{S}}_{J_0} \quad (12)$$

D. Group Method Data Handling

The idea of GMDH is to employ estimates of the output variable obtained from simple regression equations that include small subsets of input variables [23]. However, The best of these estimates are included in the set of input variables, and, again, small subsets of variables from this set are used to build new estimates [24].

The GMDH algorithm has been successfully used to deal with uncertainty and nonlinearity of systems in a wide range of application such as the signal processing, control systems, economy, ecology, medical diagnostics [25], even in chemistry [26] [27]. Some simplified approximations, which as the two directions regressive GMDH [28] and [25] investigate on the hybrid Group Method of Data Handling (GMDH) with the Wavelet Decomposition for Time Series Forecasting. The principal benefits of this hybrid GMDH are the nodal crossover through various structural layers, and instantaneous grouping of multi-input [29]

This modification and hybridization also improvement will be proposed a new hybrid GMDH that will be helped by the computational constraint. It will become more flexible as well as robust/efficient than the conventional GMDH. The GMDH is a heuristic self-organization method and a procedure for constructing a high-order polynomial of the form [30].

$$y = a_0 + \sum_i^N a_i x_i + \sum_i^N \sum_j^N a_{ij} x_i x_j + \sum_i^N \sum_j^N \sum_k^N a_{ijk} x_i x_j x_k + iNjNkNlNaijklxixjxkxl+\ldots \quad (13)$$

$$y_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 {x_i}^2 + a_5 {x_j}^2 \quad (14)$$

Where y is the output variable$x_1, x_2, \ldots, x_n$ are the input variables and a is the coefficient parameters.

III. ANALYSIS

In this paper, we use the 5-year electrical load in the holy month of Ramadan in North Sumatra, Indonesia. The first to do is perform data classification and descriptive statistics that can be seen on Table III. In general, the deviation of electrical data in 2012,2013,2014 differs from 2015. It can be identified that electricity consumption increases every year.

TABLE III
STATISTICS DESCRIPTIVE ELECTRICAL LOAD DATA

	Year			
	2012	2013	2014	2015
N	1632	1392	1440	1392
Mean	1941.8	1941.88	2164.3	2274.2
Stddev	271.29	271.29	291.28	305.02
Minimum	1329.85	1270.7	904.28	1525.8
Median	1736.5	1876.6	2110.2	2221.6
Maximum	2384.32	2644.06	2810.56	2996.50

Then, the signal is split into the mother wavelet scale. Along the way, the original signal is transformed into a signal with a lower resolution.

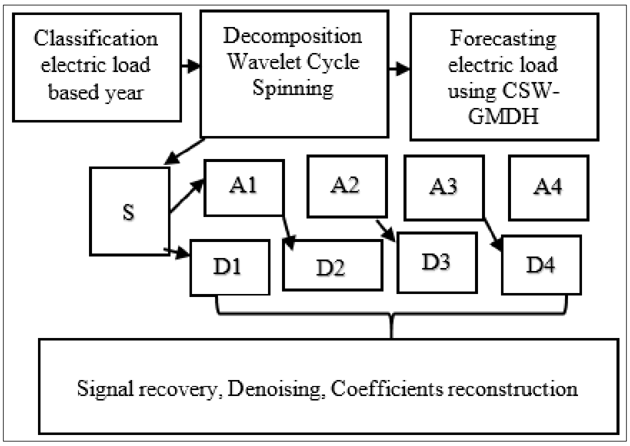

Fig. 1. Process Wavelet Decomposition.

Figure 1 describes the process of the signal is decomposed into four levels. The original signal S is decomposed into more detail (*D1*) and approaches or forecasts (*A1*). *D1* coefficient obtained from the high-pass filter, containing high-frequency components that describe the pattern of short-term. Other coefficients, *A1* is derived from the low-pass filter and contain a low-frequency component that describes the long-term pattern. If the rate of decomposition of two or higher, only coefficients approach or approximate unravel. Coefficient *Al* parsed to *A2* and *D2*. The next level of decomposition is achieved by repeating the sequence. Decomposition can continue only until the details of the individual consist of a sample or a single pixel. Once this process is completed, the signal can be smoothed by eliminating high frequency or by splitting it into low frequency and high frequency. To see the plot of the electric load in the holy month of Ramadhan can be seen in Figure 2.

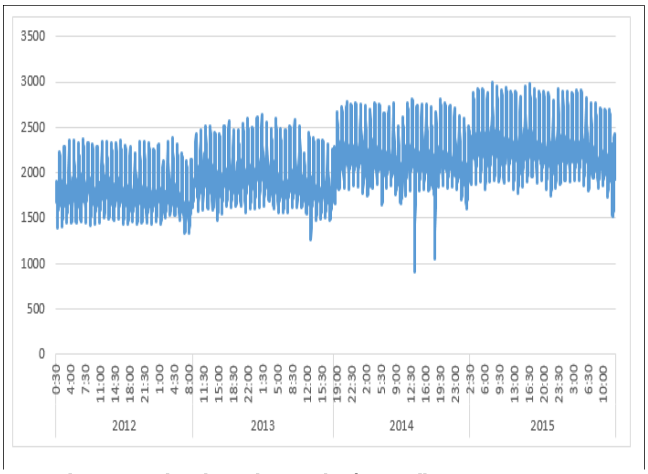

Fig. 2. Electric Load in the Holy Month of Ramadhan per 30 Minutes.

We undertake construction wavelet signal by using the data in 2015. The data from 2012 to 2014 will be used as training in wavelet GMDH, and the data in 2015 will be conducted as testing to see the validity of wavelet GMDH. Then, the data in 2016 will be compared with the result of forecasting to know the performance of CSW-GMDH. For ease of viewing characteristic of the data in 2015, we take the mean of the daily data electric load.

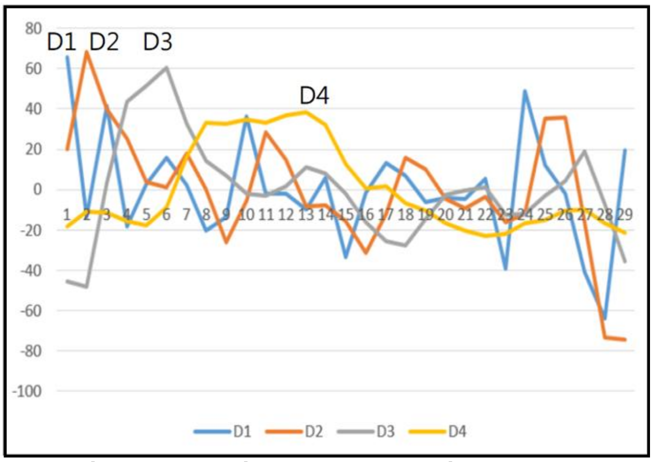

Fig. 3. Cycle Spinning Wavelet Decomposition Level 4.

For example, on the first day a holy month of Ramadhan there were 48 data of electric loads (30 minutes per day). Then, we find the mean value of these data are considered that can represent a characteristic of daily data. So, that in conducting wavelet construction only about 29 daily data holy month of Ramadhan in 2015. To see the features of construction *D* can be seen in Figure 3. After conducting testing the next step is to forecast using CSW-GMDH and compared with actual data electric load in holy Ramadhan electric load 2016 can be seen in Figure 4. Figure 5 represent the plot of testing illustrates a good to fit prediction and actual (indicated by blue and red line).

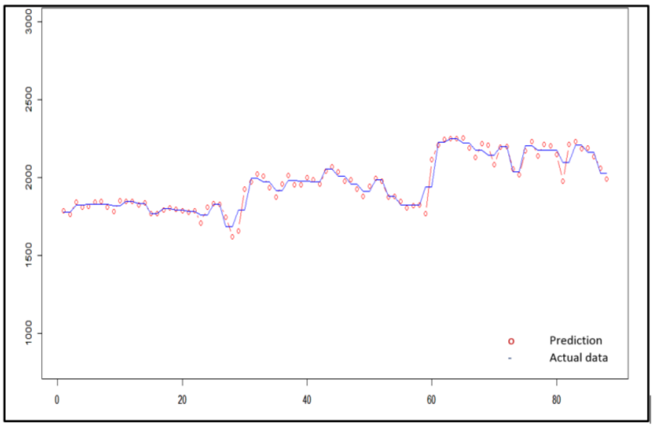

Fig. 4. Testing Cycle Spinning Wavelet GMDH.

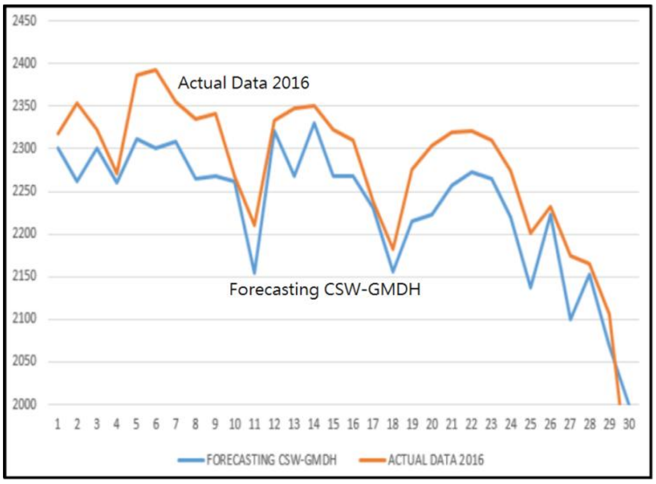

Fig. 5. Forecasting vs Actual Data.

IV. CONCLUSION

The accuracy of forecasting is fundamental to many decisions processes and hence the research for improving the effectiveness of prediction or forecasting. Based on analysis using CSW-GMDH reached the performance measures, MAE=5.50 and R^2=99.77%. It can be concluded CSW-GMDH are extremely powerful high computing. Massive parallelism makes very efficient, CSW-GMDH can learn and generalize from training data. CSW-GMDH are particularly fault tolerant. CSW-GMDH are noise tolerant. So they can cope with situations

ACKNOWLEDGMENT

We thanks to anonymous referees for their useful suggestions. This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST-107-2221-E-324-018-MY2 and partial supported by Bioinformatics and Data Science Research Center (BDSRC) Bina Nusantara University, and Department of Statistics Padjadjaran University.

REFERENCES

[1] S. Mallat, A Wavelet Tour of Signal Processing. 2009. https://www.di.ens.fr/~mallat/papiers/WaveletTourChap1-2-3.pdf

[2] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, 1989.

[3] C. Dai, D. Pi, Z. Fang, and H. Peng, "Wavelet Transform-Based Residual Modified GM(1,1) for Hemispherical Resonator Gyroscope Lifetime Prediction," IAENG International Journal of Computer Science, vol. 40, no. 4, pp. 250-256, 2013.

[4] L. W. Chew, W. C. Chia, L. minn Ang, and K. P. Seng, "An Optimum Approach for Image Compression: Tuned Degree-K Zerotree Wavelet Coding," IAENG International Journal of Computer Science, vol. 36, no.2, pp. 175–182, 2009.

[5] S. Rukun, Subanar, D. Rosadi, and Suhartono, "The Adequateness of Wavelet Based Model for Time Series," in Journal of Physics: Conference Series, pp. 1-7, 2013.

[6] S. Suhartono and S. Subanar, "Development of Model Building Procedures in Wavelet Neural Networks for Forecasting Non-stationary Time Series," European Journal of Scientific Research, vol. 34, no. 3, pp. 416-427, 2009.

[7] B. W. Otok, Suhartono, B. S. S. Ulama, and A. J. Endharta, "Design of Experiment to Optimize the Architecture of Wavelet Neural Network for Forecasting the Tourist Arrivals in Indonesia," in Communications in Computer and Information Science, vol. 253, pp. 14-23, 2011.

[8] B. Warsito, Subanar, and Abdurakhman, "Wavelet Decomposition for Time Series: Determining Input Model by using MRMR criterion," Hacettepe Journal of Mathematics and Statistics, vol. 43, no. 3, pp. 1-12, 2015.

[9] U. Sa'Adah, Subanar, S. Guritno, and Suhartono, "Wavelet Neural Network Model Selection for Nonlinear-Seasonal Time Series Forecasting," Global Journal of Pure and Applied Mathematics, vol. 11, no. 1, pp. 25-36, 2015.

[10] R. E. Caraka, S. A. Bakar, and G. Anuraga, "The Step Construction of Penalized Spline in Electrical Power Load Data," Telkomnika, vol. 17, no. 2, pp. 1023-1031, 2019.

[11] R. E. Caraka and S. A. Bakar, "Evaluation Performance of Hybrid Localized Multi Kernel SVR (LMKSVR) In Electrical Load Data Using 4 Different Optimizations," Journal of Engineering and Applied Science, vol. 13, no. 17, 2018.

[12] Y. F. Ju and S. W. Wu, "Village Electrical Load Prediction by Genetic Algorithm and SVR," in Proceedings - 2010 3rd IEEE International Conference on Computer Science and Information Technology, ICCSIT, pp. 278-281, 2010.

[13] Suparti, R. . Caraka, B. Warsito, and H. Yasin, "The Shift Invariant Discrete Wavelet Transform (SIDWT) with Inflation Time Series Application," Journal of Mathematics Research, vol. 8, no. 4, pp. 14-20, 2016.

[14] A. Haar, "Zur Theorie der Orthogonalen Funktionensysteme," Mathematische Annalen, vol. 71, no. 1, pp. 38-53, 1911.

[15] M. J. Shensa, "The Discrete Wavelet Transform: Wedding the À Trous and Mallat Algorithms," IEEE Transactions on Signal Processing, vol. 40, no. 10, pp. 2464–2482, 1992.

[16] L. Chun-Lin, "A Tutorial of the Wavelet Transform," Microb. Fuel Cells, pp. 2–71, 2008.

[17] O. Christensen, "An Introduction to Wavelet Analysis," in Applied and Numerical Harmonic Analysis, 2010.

[18] G. P. Nason, "Wavelet Shrinkage Using Cross-Validation," Journal of The Royal Statistical Society Series B-Statistical Methodology, 2018.

[19] R. R. Coifman, Y. Meyer, and M. V Wickerhauser, "Wavelet Analysis and Signal Processing," in Wavelets and Their Applications, pp. 153–178, 1992.

[20] N. Ul Hasan, W. Ejaz, M. K. Atiq, and H. S. Kim, "Recursive Pyramid Algorithm-based Discrete Wavelet Transform for Reactive Power Measurement in Smart Meters," Energies, vol. 6, pp. 4721–4738, 2013.

[21] M. Vishwanath, "The Recursive Pyramid Algorithm for the Discrete Wavelet Transform," IEEE Transactions on Signal Processing, vol. 42, no. 3, pp. 673–676, 1994.

[22] Suhartono, B. S. S. Ulama, and A. J. Endharta, "Seasonal Time Series Data Forecasting by using Neural Networks Multiscale Autoregressive Model," The American Journal of Applied Sciences, vol. 7, no. 10, pp. 1372–1378, 2010.

[23] J.-A. Mueller and F. Lemke, Self-organising Data Mining: An Intelligent Approach to Extract Knowledge from Data. 1999.

[24] M. C. Acock and Y. A. Pachepsky, "Estimating Missing Weather Data for Agricultural Simulations Using Group Method of Data Handling," Journal of Applied Meteorology and Climatology, vol. 39, pp. 1176–1184, 2002.

[25] M. S. Voss and X. Feng, "A New Methodology For Emergent System Identification Using Particle Swarm Optimization ({PSO}) And The Group Mehtod Data Handling ({GMDH})," in GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 1227–1232, 2002.

[26] A. Rostami, A. Hemmati-Sarapardeh, A. Karkevandi-Talkhooncheh, M. M. Husein, S. Shamshirband, and T. Rabczuk, "Modeling Heat Capacity of Ionic Liquids Using Group Method of Data Handling: A Hybrid and Structure-based Approach," International Journal of Heat and Mass Transfer, vol. 129, pp. 7–17, 2019.

[27] S. Atashrouz, G. Pazuki, and S. S. Kakhki, "A GMDH-type Neural Network for Prediction of Water Activity in Glycol and Poly (Ethylene Glycol) Solutions," Journal of Molecular Liquids, vol. 202, pp. 95–100, 2015.

[28] X. Wang, L. Li, D. Lockington, D. Pullar, and D.-S. Jeng, "Self-Organizing Polynomial Neural Network for Modelling Complex Hydrological Processes," Research Report - University of Sydney, Department of Civil Engineering, no. 861, pp. 1-29, 2005.

[29] S. Atashrouz, G. Pazuki, and Y. Alimoradi, "Estimation of the Viscosity of Nine Nanofluids using a Hybrid GMDH-type Neural Network System," Fluid Phase Equilib., vol. 372, pp. 43–38, 2014.

[30] A. G. Ivakhnenko, "The Group Method of Data of Handling; A Rival of the Method of Stochastic Approximation," Sov. Autom. Control, vol. 13, pp. 43–55, 1968.

Rezzy Eko Caraka received The B.S. degree (S.Si) from Department of Statistics Diponegoro University in 2015 and Master of Science by research (MSc-Res) School of Mathematical Sciences, FST, the National University of Malaysia. In 2019 He starts Doctor of Philosophy, College of Informatics, Chaoyang University of Technology, Taiwan. He is active as researcher in Bioinformatics & Data Science Research Center University of Bina Nusantara (BINUS) and Department of Statistics, Padjadjaran University. He was co-founder Statistical Calculator (STATCAL).

Rung-Ching Chen received a B.S. from the Department of Electrical Engineering in 1987, and an M. S. from the Institute of Computer Engineering in 1990, both from National Taiwan University of Science and Technology, Taipei, Taiwan. In 1998, he received his Ph.D. from the Department of Applied Mathematics in computer science, National Chung Hsing University. He is now a distinguished professor in the Department of Information Management, Taichung, Taiwan. His research interests include network technology, pattern recognition, and knowledge engineering, IoT and data analysis, and applications of Artificial Intelligence.

Toni Toharudin currently Associate Professor at the Department of Statistics, Padjadjaran University. Toni does research in Statistics. He received the Master of Science University of Leuven Belgium and Ph.D. Spatial Sciences University of Groningen,Moreover, Toni act as head of research group in time series and regression.

Bens Pardamean He earned a doctoral degree in informative research from the University of Southern California in 2007, as well as a master's degree in computer education and a bachelor's degree in computer science from California State University, Los Angeles in 1988 and 1994, respectively. He currently holds a dual-appointment as the Director of Bioinformatics & Data Science Research Center and as an Associate Professor of Computer Science at the University of Bina Nusantara (BINUS) in Jakarta, Indonesia.

Sakhinah Abu Bakar currently works at School of Mathematical Sciences, The National University of Malaysia. She earned Master of Science in Mathematics from Universiti Teknologi Malaysia (UTM) and her PhD in Computer Science from The University of Sydney, Australia. Her research interest including Computational Mathematics, Intelligent Systems, Machine learning and Network Science.

Hasbi Yasin received a B.S (S.Si) Department of Mathematics, Diponegoro University in 2005, and Master of Science (M.Si) Department of Statistics, ITS, Surabaya in 2009. He is now Assistant Professor in Department of statistics Diponegoro University. His research interests include spatial statistics, computational statistics.