An Approach of Correlation Inter-variable Modeling with Limited Data for Inter-Bus Transformer Weather Sensitive Loading Prediction

Fitriana[1] and Nanang Hariyanto[2]

[1]Research Center for Electrical Engineering and Mechatronics Indonesian Institute of Sciences
Komplek LIPI Gedung 20 Jalan Cisitu 21/154 D Dago Bandung 40135 Indonesia
[2]School of Electrical Engineering and Informatics, Institut Teknologi Bandung
Gedung Kerma PLN-ITB Lt. 2 Jalan Ganesha No.10 Bandung 40132 Indonesia
[1]fitrisalsa@gmail.com, [2]nanang.hariyanto@stei.itb.ac.id

Abstract: This paper deals with a method to determine short-term loading estimation of inter-bus transformers based only a limited data of variables. A grey predicting approach is implemented in this paper to solve dynamic short term load forecasting affected by the climate changes in Indonesian EHV grid, as provided by Jawa Bali Control Center. A dynamic forecasting model is needed due to the uncertain nature of the load predicting process specifically when load changes is correlated with the external effect. Traditional GM (1,1) as basic model is presented to compare with the correlation inter-variable model in grey method for inter-bus loading transformers loading conditions refer to the temperature variation pattern during only the year 2009. GM(1,2) model denotes the correlation and relationship of hourly load demand affected to the 2 (two) tropical seasons, in this case by using local ambient temperature as dynamic variable. The daily load curve for Cawang and Cibatu distribution areas, as representing big local substations in Indonesia spatially, was used to validate the models. The model adequacy and weekly forecasting result had indicated a good and justified grade in error diagnostic checking (MAPE) in each location.

Keywords: grey dynamic forecasting model, correlation inter-variable model, ambient temperature records, local load profiles

1. Introduction

Rapid information and development of new theories of systems science have become an important part of modern science and technology, especially in research of uncertainty system. There are several theories how to handle uncertain data: fuzzy logic, system identification, dimensional analysis and Grey System Theory (GST). GST is the one of the most important research of uncertainty system that realizes the correct description and effective supervision in operation action and evolution law of system, mainly through generating, exploring and extracting valuable information from a little part known information to predict the unknown information value, where appeared initially in early 1980 by J. Deng [1].

In grey systems, the information have partially known and partially unknown parameters especially useful, (1) when the complete set of factors involved in the system's behaviour is unclear, (2) when the relationships of system factors to the system's behaviour and inter-relationships among factors are uncertain, (3) when the system behaviour is too complex to determine completely, or (4) when only limited information or time series data on system behaviour is available. The essential scopes of GST encompass in research methods about: grey relational space, grey generating space, grey forecasting, grey decision making, grey control, grey mathematics and grey theory as base [1].

In real system, daily load requirement would change in real time due to uncertainty load behaviour in electric power systems. National State Electricity Company in Indonesia (PLN) have developed load forecasting based on coefficient method to support its operation and

Received: June 11th, 2012. Accepted: December 29th, 2012

control planning [2]. It have been implemented only for a high and macro-scale data in Jawa Bali load systems using minimum 3-5 years past data. Unfortunately, its method do not accommodate the real relationship between daily loads and many weather variables such as ambient temperature, relative humidity, wind speed, etc. Regarding the changing climate in tropical area Indonesia, the electricity demand also should be predicted by extrapolating a predetermined relationship between the load and its influential variables - namely, time and/or weather-sensitive, toward short-term or long-term load forecasting. This paper propose a model scheme for quantifying the relationship between the load and ambient temperature data information specific in certain area and the use of a suitable parameter estimation technique obtained from Grey Forecasting Method (GFM) in accordance with the schemes. However, due having only a little temperature data and smaller case coverage, only two main schemes of the GFM as grey dynamic models, abbreviated as GM (1,1) and GM (1,2), will be implemented [3]. It was used to predict of short term electric demand that includes certain location (as condition referred to 'where') as one of its main elements, in addition to load magnitude (how much) and seasonal or temporal (when) characteristics, so the method have been implemented as spatial load forecasting [9].

As study cases, 2 (two) inter-bus transformers loading conditions in Jakarta area had been chosen to improve the grey predicting approach in power/energy systems. The application GM(1,2) combined model considering the influencing factor of load to forecast a day-ahead electricity prices had been implemented in China [6], and many additional methods also can be adopted to obtain optimal coefficient and enhance the estimation validation index [4],[5]. However, new contribution in this paper was grey theory implementation in loading estimation analysis with incomplete data and correlate 2 (two) inter-variables, where known the specific condition in one area which are slightly different with another, inclusive weather effects, power consumption pattern of its customer, sudden network change due to operation strategies requirement. The conformity of correlation coefficient based on the profiling of data information has developed a grey model scheme appropriate in specific area representing tropical climate condition. Hence, the hybrid grey model developed by [5] could not be applied in tropical climate condition at all, so the GM(1,2) method schemes mentioned in [4] did not appropriate to all cases.

2. Grey Estimation Methods

In grey system theory for analysing non-linier systems, a differential model called the Grey Model (GM) only need to use at least 4 (four) samples to model and replace difference modelling in vast quantities of data [3].

A. GM(1,1) Model

The series $X^{(0)}$ processed is ordered to be the GM (1,1) modelling sequence, denote the original data sequence as follows:

$$X^{(0)} = \left\{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3).....,X^{(0)}(n)\right\} \tag{1}$$

The AGO formation of $X^{(0)}(k) \geq 0, k = 1,2,3,......, n)$ is generated to the first order series $\left\{X^{(1)}(k)\right\}, \left\{X^{(1)}(k)\right\}$ is defined as follows:

$$X^{(1)} = \left\{X^{(1)}(1), X^{(1)}(2), X^{(1)}(3),....., X^{(1)}(n)\right\} \tag{2}$$

The above two series meet the following relationship as follows:

$X^{(1)}(1) = X^{(0)}(1)$, and

$$\underset{Grey}{\overset{G}{\uparrow}} \quad \underset{Model}{\overset{M}{\uparrow}} \quad \underset{1st\ Order}{\overset{(1,}{\uparrow}} \quad \underset{One\ Variable}{\overset{1)}{\uparrow}}$$

With GM (1,1) model for demand forecasting, but the accuracy of forecasting model is also need to be checked so that the future value predicted have a higher credibility. Residual test as follows: $\varepsilon^{(0)}(k) = X^{(0)}(k) - \hat{X}^{(0)}(k)$ with absolute error in percentage (MAPE) and $\Delta = \varepsilon^{(0)}(k) / X^{(0)}(k)$ [7].

B. GM(1,2) Model

One kind of the grey relationship model in grey theory is GM(1,N), where the data can be separated into two sequences, i.e. one major sequence factor, which is the sequence that masters the systems represent such as equation (1); and N-1 influencing sequence factors, which are the sequences that influence the systems. The major sequence factor and its sequence factors can be defined as follows:

$$\begin{aligned} X_2^{(0)} &= \left\{X_2^{(0)}(1), X_2^{(0)}(2), X_2^{(0)}(3), \ldots\ldots, X_2^{(0)}(n)\right\} \\ X_3^{(0)} &= \left\{X_3^{(0)}(1), X_3^{(0)}(2), X_3^{(0)}(3), \ldots\ldots, X_3^{(0)}(n)\right\} \\ &\vdots \\ X_N^{(0)} &= \left\{X_N^{(0)}(1), X_N^{(0)}(2), X_N^{(0)}(3), \ldots\ldots, X_N^{(0)}(n)\right\} \end{aligned} \tag{11}$$

where N is defined as the original sequence number. Subjecting these N-1 sequences to AGO as in (3) obtain the sequences $X_N^{(1)}$.

After these sequences are subjected to the Accumulating Generation Operation (AGO), refer to equation (3), the following sequences $X_N^{(1)}$ obtained. The grey differential equation of GM(1,N) model is [3]

$$\frac{dX^{(1)}(k)}{dt} + aX_1^{(1)}(k) = \sum_{i=2}^{N} b_i X_i^{(1)}(k) \tag{12}$$

where a is the develop factor and b_i as the driving terms which is the relationship weighting factors. So, the parameter vector defined in equation (6) can be solved where:

$$\hat{a} = \left[a, b_1, b_2, \ldots., b_N\right]^T = (B^T \bullet B)^{-1} \bullet B^T \bullet Y \tag{13}$$

Which, B is a matrix derived from the matrix equation representing the grey equation in (12).

$$B = \begin{bmatrix} -\frac{1}{2}\left[X^{(1)}(1) + X^{(1)}(2)\right] & X_2^{(1)}(2) & \cdots & X_N^{(1)}(2) \\ -\frac{1}{2}\left[X^{(1)}(2) + X^{(1)}(3)\right] & X_2^{(1)}(3) & & X_N^{(1)}(3) \\ \vdots & & & \vdots \\ -\frac{1}{2}\left[X^{(1)}(n-1) + X^{(1)}(n)\right] & X_2^{(1)}(n) & & X_N^{(1)}(n) \end{bmatrix} \tag{14}$$

$$\hat{x}_1^{(1)}(k) = \left[x_1^{(0)}(1) - \frac{b}{a}x_2^{(1)}(k)\right]e^{-a(k-1)} + \frac{b}{a}x_2^{(1)}(k) \tag{19}$$

where $X_1^{(1)}(0)$ is taken to be $X_1^{(0)}(1)$. Finally the restoration of Inverse AGO will use equation (20) below.

$$\hat{X}^{(0)}(k+1) = \hat{X}^{(1)}(k+1) - \hat{X}^{(1)}(k) \tag{20}$$

3. The Profiling of Load Models

The sample data in this study are based on the historical loading of inter-bus transformers (IBT) and ambient temperature information in 2 (two) local substations in Jakarta area, i.e.: Cawang substation in 2009, only data until September 28, 2009 when IBT fail loading in there, and Cibatu substation with 2009 data records from January 1 until December 31, 2009.

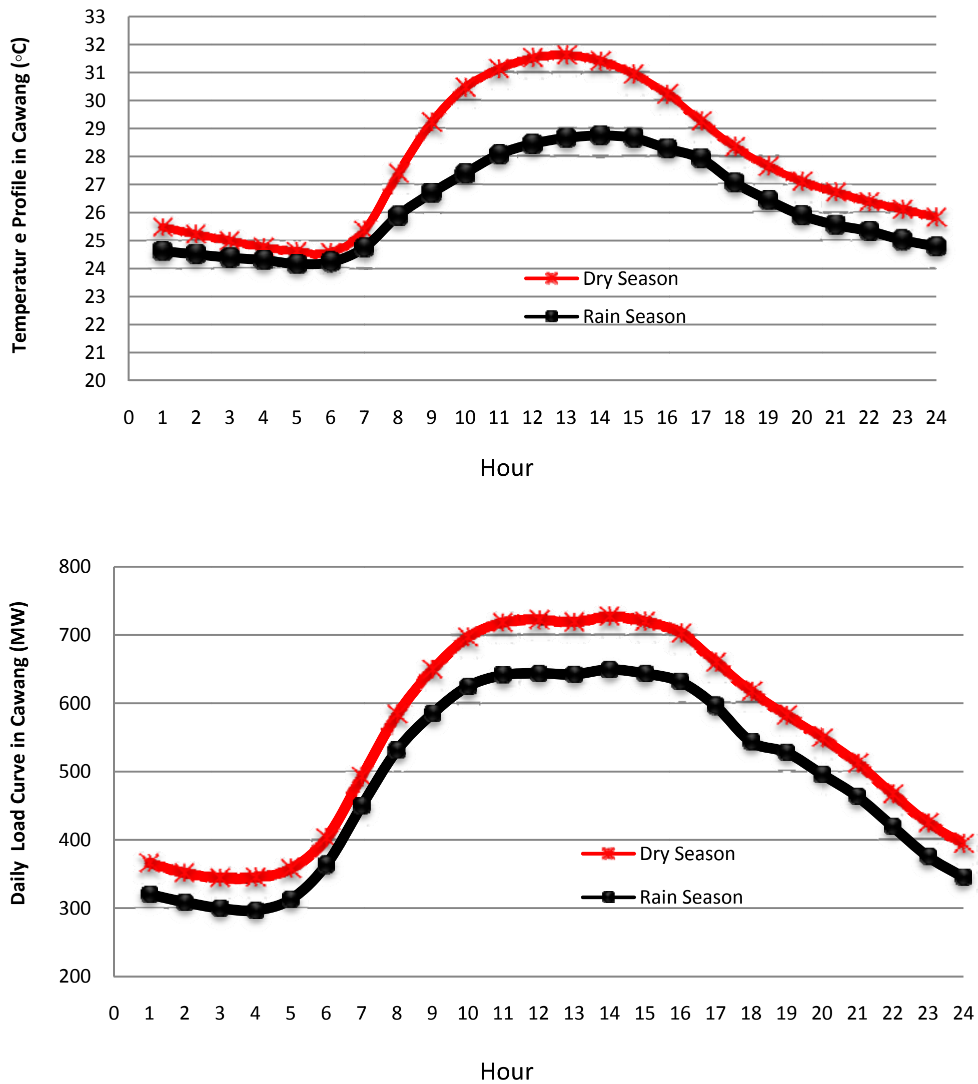

Figure 2. Profile of Weather-Sensitive Daily Loads in Cawang

After profiling data, it was found that daily load curve of each day in Cawang in one week not be similar, as well as in Cibatu. But ambient temperature patterns always similar in a week for same season. The clustering tendency of daily load distribution in specific location should be compared with the trend of ambient temperature above normal, and then a correlation between load increments due to seasonal changes was represented by the increment of ambient temperature during the dry season. From Figure 2, the effect of seasonal changes have been faced on the average loading increase in that location, so the linear correlation states affected temperature and load model can be used to determine the sensitivity of IBT loading.

The abnormal loading conditions in Cawang in 2009 have been described using box-plot to find the monthly loading distribution. It was used as boundary of the data sampling that represent each IBT loading condition in certain season, as described in Figure 3. The daily load pattern curve in Cawang dominated by business customers and commercial linkages shows the correlation between the load and ambient temperature in site. As boundary, correlation modelling in Cawang for the dry season occurs from March to September and for rainy season occurs from January to February.

Using the same profiling method, the temperature affected by IBT loading in Cibatu was determined and form weather sensitive load modeling pattern which the dry season occurs from March to October. Overall the load increase pattern due to the temperature effect in Cawang occurs starting at hour 7 to 18, who dominated the business customer activity (as representation of the commercial load profile), while in Cibatu suggests the contribution of industrial customers (represented the industry load profile) with mid-day load fluctuation.

4. Proposed Model Scheme Based on Correlation Inter-variable

In this paper, daily day-ahead load in MW distributed in Cawang and Cibatu are selected to forecast and validate the performance of the common Grey estimation model. i.e. GM(1,1) and GM(1,2). Each model was applied to the same hours and the same days matched with the particular season to find the output coefficients therein.

In general, variables that act upon the system of interest should be external or predefined. The GM(1,1) model is a single sequence modelling, which makes use of only the system's behaviour sequence, represented by output sequence or background values, without considering any external acting sequences represented by some input sequences or driving quantities as well as in GM(1,N) model [3].

The parameters (-a) and (b) mentioned in (9) and (19) are called the development coefficient and grey action quantity, respectively. The parameter (-a) reflects development states of $\hat{X}^{(1)}$ and $\hat{X}^{(0)}$ and the grey action quantity in GM(1,1) is a value derived from the background values contained in column 1 of matrix B in (6). Interestingly in GM(1,N) model, (-a) still represent the development coefficient of the system, but $b_i X_i^{(1)}(k)$ is called the driving term which (b_i) reflect the driving coefficients. It would be different among (b) in GM(1,1) with (b_i) in GM(1,2) which had been used in this paper. All b_i parameters derived from simulation in each season was used to compare the impact correlations of daily weather trend and daily load curve in different areas, where shown in Table 1.

Table 1. Correlation Coefficient Comparison in Cawang

SEASON	Driving Parameter or Model Control Variable (b_i)						
	MON	TUE	WED	THU	FRI	SAT	SUN
DRY	39.8	36.1	40.3	37.5	35.2	31.9	30.0
RAIN	37.3	37.8	33.9	31.4	35.3	26.2	25.1

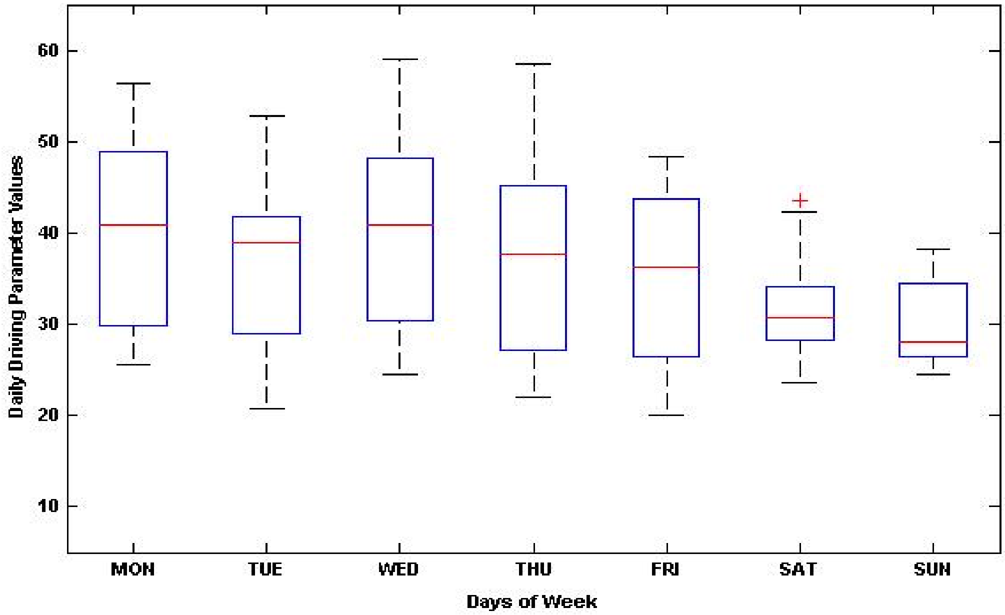

Figure 5. Correlation Coefficient in GM(1,2) for All Days of Week through Dry Season

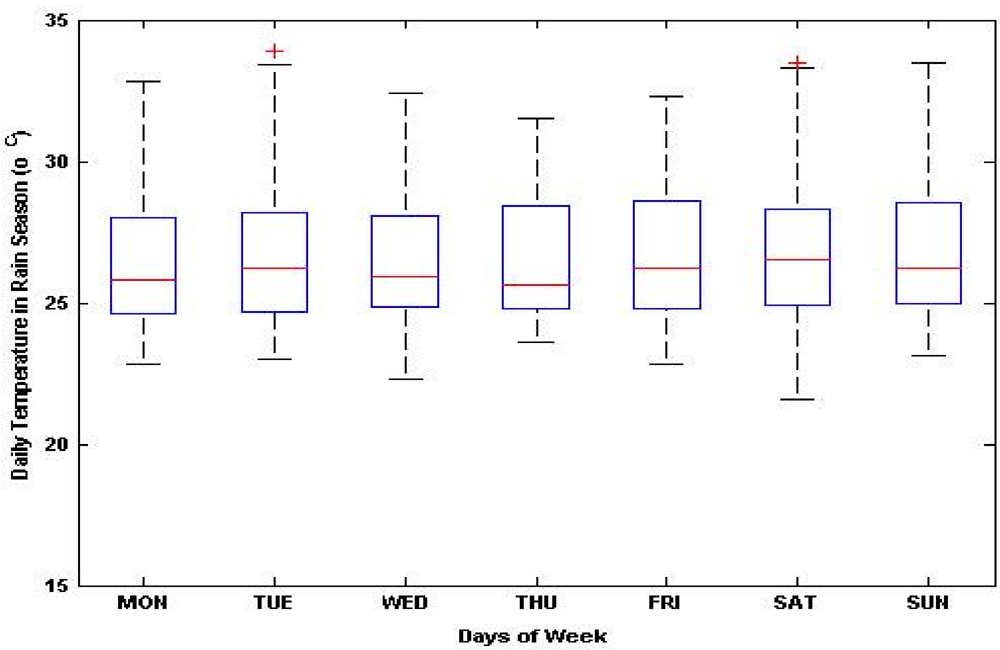

Figure 6. Ambient Temperature Distribution through Rainy Season

Hence, it can be seen that implementation GM(1,2) only in workday period has lower error rather than the model scheme in all day consistently, likewise the model precision in Cawang has better accuracy compared to Cibatu. The adequacy of proposed Grey model were tested through the diagnostic checking of residuals, with MAPE values were below 10%, which were also included justified grade in model adequacy limit, as defined in [3].

Table 3. Model Error Comparison in Dry Season Profiling

MAPE (%)	GM(1,2)		Proposed Model Scheme	
	CAWANG	CIBATU	CAWANG	CIBATU
MONDAY	4.06	7.35	4.06	7.35
TUESDAY	4.20	5.87	4.20	5.87
WEDNESDAY	4.52	7.39	4.52	7.39
THURSDAY	4.92	8.93	4.92	8.93
FRIDAY	5.39	9.17	5.39	9.17
SATURDAY	5.55	13.63	2.57	4.86
SUNDAY	5.26	17.49	2.34	7.95
AVERAGE ERROR	4.84	9.98	**4.00**	**7.36**

Table 4. Model Error Comparison in Rainy Season Profiling

MAPE (%)	GM(1,2)		Proposed Model Scheme	
	CAWANG	CIBATU	CAWANG	CIBATU
MONDAY	6.58	7.54	6.58	7.54
TUESDAY	5.63	6.91	5.63	6.91
WEDNESDAY	7.04	6.56	7.04	6.56
THURSDAY	7.64	8.77	7.64	8.77
FRIDAY	6.01	7.98	6.01	7.98
SATURDAY	8.10	9.88	1.73	4.92
SUNDAY	7.51	11.83	4.96	4.50
AVERAGE ERROR	6.93	8.49	**5.66**	**6.74**

For Jakarta area spatially, the proposed grey model scheme was used to predict one-hour ahead of load value in each location during the same season. From equation (9) and (19), there were 24 hours x 7 days x 4 variable values that would produce 24 sequence sets which are capable for establishing 24 GM(1,2) models to predict one-workday ahead daily loads and 24 GM(1,1) models to predict one-weekday ahead daily load, respectively in a week. In workday forecasting, one temperature value must be included to increase the last number of AGO simulated, $X_2^{(1)}(k)$ as initial value in (19).